U.S. Application No. 10/550,797
For IDS filed May 1, 2007

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, 14/47, 16/18, 16/28, C12Q 1/68, A61K 38/17

A2

(11) International Publication Number: **WO 00/12711**

(43) International Publication Date: 9 March 2000 (09.03.00)

(21) International Application Number: PCT/US99/20468

(22) International Filing Date: 2 September 1999 (02.09.99)

(30) Priority Data:

09/145,815	2 September 1998 (02.09.98)	US
Not furnished	2 September 1998 (02.09.98)	US
09/191,283	12 November 1998 (12.11.98)	US
Not furnished	12 November 1998 (12.11.98)	US
09/208,821	9 December 1998 (09.12.98)	US
Not furnished	9 December 1998 (09.12.98)	US
09/237,506	26 January 1999 (26.01.99)	US
Not furnished	26 January 1999 (26.01.99)	US
09/247,891	10 February 1999 (10.02.99)	US
Not furnished	10 February 1999 (10.02.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	Not furnished (CIP)
Filed on	2 September 1998 (02.09.98)
US	09/191,283 (CIP)
Filed on	12 November 1998 (12.11.98)
US	Not furnished (CIP)
Filed on	12 November 1998 (12.11.98)
US	09/145,815 (CIP)
Filed on	2 September 1998 (02.09.98)
US	09/208,821 (CIP)
Filed on	9 December 1998 (09.12.98)
US	Not furnished (CIP)
Filed on	9 December 1998 (09.12.98)
US	09/237,506 (CIP)
Filed on	26 January 1999 (26.01.99)
US	Not furnished (CIP)
Filed on	26 January 1999 (26.01.99)
US	09/247,891 (CIP)
Filed on	10 February 1999 (10.02.99)
US	Not furnished (CIP)
Filed on	10 February 1999 (10.02.99)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

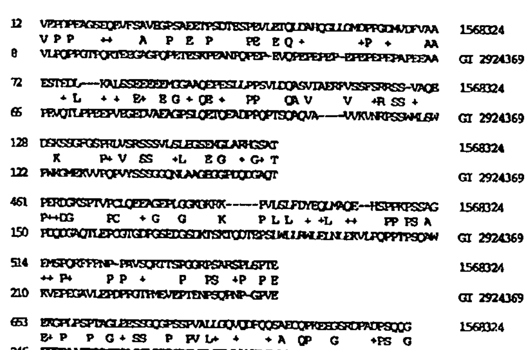
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: AU-YOUNG, Janice [US/US]; 1419 Kains Avenue, Berkeley, CA 94702 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). REDDY, Roopa [IN/US]; 1233 W. McKinley Drive, Sunnyvale, CA 94086 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue, #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). GORGONE, Gina [US/US]; 1253 Pinecrest Drive, Boulder Creek, CA 95006 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: HUMAN MEMBRANE CHANNEL PROTEINS



(57) Abstract

The invention provides new human membrane channel proteins (MECHP) and polynucleotides which identify and encode MECHP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of MECHP.

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave	TM	Turkménistan
BF	Burkina Faso	GR	Grèce		de Macédoine	TR	Turquie
BG	Bulgarie	HU	Hongrie	ML	Mali	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MN	Mongolie	UA	Ukraine
BR	Brésil	IL	Israël	MR	Mauritanie	UG	Ouganda
BY	Bélarus	IS	Islande	MW	Malawi	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	MX	Mexique	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Pays-Bas	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NO	Norvège	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire	NZ	Nouvelle-Zélande		
CM	Cameroun		démocratique de Corée	PL	Pologne		
CN	Chine	KR	République de Corée	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Roumanie		
CZ	République tchèque	LC	Sainte-Lucie	RU	Fédération de Russie		
DE	Allemagne	LI	Liechtenstein	SD	Soudan		
DK	Danemark	LK	Sri Lanka	SE	Suède		
EE	Estonie	LR	Libéria	SG	Singapour		

HUMAN MEMBRANE CHANNEL PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human membrane channel proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, immune/inflammatory, transport/secretory, osmoregulatory, muscular, cardiovascular, and neurological disorders.

BACKGROUND OF THE INVENTION

Channel proteins facilitate the transport of hydrophilic molecules across membranes by forming aqueous pores that can perforate a lipid bilayer. Many channels consist of protein complexes formed by the assembly of multiple subunits, at least one of which is an integral membrane protein that contributes to formation of the pore. In some cases, the pore is constructed to selectively allow passage of only one or a few molecular species. Distinct types of membrane channels that differ greatly in their distribution and selectivity include: (1) aquaporins, which transport water; (2) protein-conducting channels, which transport proteins across the endoplasmic reticulum membrane; (3) gap junctions, which facilitate diffusion of ions and small organic molecules between neighboring cells; and (4) ion channels, which regulate ion flux through various membranes.

Aquaporins

Aquaporins (AQP) are channels that transport water and, in some cases, nonionic small solutes such as urea and glycerol. Water movement is important for a number of physiological processes including renal fluid filtration, aqueous humor generation in the eye, cerebrospinal fluid production in the brain, and appropriate hydration of the lung. A variety of aquaporins have been found in higher animals, plants and microorganisms. The mammalian aquaporins appear to have selective expression in particular tissues, with AQP0 localized to lens epithelium; AQP1 localized to many tissues including red blood cells, kidney, eye, lung, choroid plexus, bile duct, and vascular epithelium; AQP2 localized to the apical membrane of kidney collecting duct cells; AQP3 localized to kidney, colon, trachea, urinary bladder, skin, and sclera of eye; AQP4 localized to kidney, colon, trachea, stomach, skeletal muscle, spinal cord, brain, and retina; AQP5 localized to the apical membranes of exocrine tissues; AQP6 localized to kidney; and AQP7 localized to testis (King, L.S. and P. Agre (1996) Annu. Rev. Physiol. 58:619-648; Ishibashi, K. et al. (1997) J. Biol. Chem. 272:20782-20786). AQP9 is expressed in peripheral leukocytes, less abundantly in liver, even less in lung and spleen, and not at all in thymus tissue (Ishibashi, K. et al. (1998) Biochem.

Biophys. Res. Commun. 244:268-274).

Aquaporins are members of the major intrinsic protein (MIP) family of membrane transporters. MIP family proteins are composed of four subunits, each of which may span the membrane six times, and have their N-and C-termini facing the cell cytoplasm. Proteins from bacteria, yeast, plants, and animals have been shown to be members of the MIP family (Reizer, J. et al. (1993) Crit. Rev. Biochem. 28:235-257). Aquaporin subunits are integral membrane proteins with six transmembrane regions and two conserved Asn-Pro-Ala (NPA) boxes (which are sometimes found as Asn-Pro-Ser) found in loop regions between the transmembrane regions (King, supra; Ishibashi, (1997) supra). The study of aquaporins may have relevance to understanding edema formation and fluid balance in both normal physiological and disease states (King, supra). Mutations in AQP2 cause autosomal recessive nephrogenic diabetes insipidus (Online Mendelian Inheritance in Man (OMIM) *107777 Aquaporin 2; AQP2). Reduced AQP4 expression in skeletal muscle may be associated with Duchenne muscular dystrophy (Frigeri, A. et al. (1998) J. Clin. Invest. 102:695-703). Mutations in AQP0 cause autosomal dominant cataracts in mice (OMIM *154050 Major Intrinsic Protein of Lens Fiber; MIP).

Protein-Conducting Channels

Secreted and integral membrane proteins are transported from the cytoplasm to the endoplasmic reticulum (ER) through protein-conducting channels in the ER membrane. The channel is used for both co- and post-translational translocation. In the co-translational process, transport is initiated by the action of a cytoplasmic signal recognition particle (SRP) which recognizes a signal sequence on a growing, nascent polypeptide and binds the polypeptide and its ribosome complex to the ER membrane through an SRP receptor located on the membrane. The ribosome complex, together with the attached polypeptide, becomes membrane bound. As the nascent chain emerges from the ribosome, it is fed into the channel and across the ER membrane. The post-translational process also requires a signal sequence on the protein to be translocated, but does not require an SRP. The protein enters the channel and is driven across the ER membrane by the hydrolysis of adenosine triphosphate (ATP) by BiP, an ATPase and molecular chaperone in the ER lumen.

The protein-conducting channel, termed the Sec61p complex, is composed of multiple, probably two, heterotrimers of three membrane proteins, the alpha, beta, and gamma subunits of Sec61p. The Sec61p complex forms a ring structure visible by electron microscopy (EM). EM and quenching experiments indicate a channel diameter of 20 to 60 Å. Association of the Sec61p complex with the ribosome and with the proteins Sec62p, Sec63p, Sec71p, Sec72p, BiP, and TRAM (translocating chain-associating membrane protein) is required for some of the channel's

functions. The Sec61p alpha subunit contains ten membrane-spanning segments and has been found to line the path of the translocating polypeptide chain from one side of the membrane to the other. The sequences of dog and rat Sec61p alpha genes have been determined. Homologs of the mammalian Sec61p alpha are found in the yeast Saccharomyces cerevisiae (Sec61p) and in bacteria (SecYp). (See Görlich, D. et al. (1992) Cell 71:489-503; Matlack, K.E.S. et al. (1998) Cell 92:381-390.)

Defects in protein trafficking to organelles or to the cell surface are involved in numerous human diseases and disorders including cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease. Cancer cells secrete excessive amounts of hormones or other biologically active peptides.

Gap Junctions

Gap junctions (also called connexons) are channels that function chemically and electrically to couple the cytoplasms of neighboring cells in many tissues. Gap junctions function as electrical synapses for intercellular propagation of action potentials in excitable tissues. In nonexcitable tissues, gap junctions have roles in tissue homeostasis, coordinated physiological response, metabolic cooperation, growth control, and the regulation of development and differentiation. Gap junctions help to synchronize heart and smooth muscle contraction, speed neural transmission, and propagate extracellular signals. Gap junctions can open and close in response to particular stimuli (e.g., pH, Ca^{+2}, and cAMP). The effective pore size of a gap junction is approximately 1.5 nm, which enables small molecules (e.g., those under 1000 daltons) to diffuse freely through the pore. Transported molecules include ions, small metabolites, and second messengers (e.g., Ca^{+2} and cAMP).

Each connexon is composed of six identical subunits called connexins. At least thirteen distinct connexin proteins exist, with each having similar structures but differing tissue distributions. Structurally, the connexins are integral membrane proteins with four putative membrane spanning regions and N- and C-termini oriented towards the cell cytoplasm. Conserved regions include the membrane spanning regions and two extracellular loops. The variable regions, which are two cytoplasmic loops and the C-terminal region, may be responsible for the regulation of different connexins. (See Hennemann, H. et al. (1992) J. Biol. Chem. 267:17225-17233; PRINTS PR00206 connexin signature.)

Connexins have many disease associations. Female mice lacking connexin 37 (Cx37) are infertile due to the absence of the oocyte-granulosa cell signaling pathway. Mice lacking Cx43 die shortly after birth and show cardiac defects reminiscent of some forms of stenosis of the

pulmonary artery in humans. Mutations in Cx32 are associated with the X-linked form of Charcot-Marie-Tooth disease, a motor and sensory neuropathy of the peripheral nervous system. Cx26 is expressed in the placenta, and Cx26-deficient mice show decreased transplacental transport of a glucose analog from the maternal to the fetal circulation. In humans, Cx26 has been identified as the first susceptibility gene for non-syndromic sensorineural autosomal deafness. Cx46 is expressed in lens fiber cells, and Cx46-deficient mice develop early-onset cataracts that resemble human nuclear cataracts. (See Nicholson, S.M. and R. Bruzzone (1997) Curr. Biol. 7:R340-R344.)

Ion Channels

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form an ion-selective pore within the membrane. There are two basic types of ion channel: ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

Ion channels share common structural and mechanistic themes. The channel consists of four or five subunits or protein monomers that are arranged like a barrel in the plasma membrane. Each subunit typically constists of six potential transmembrane segments (S1, S2, S3, S4, S5, and S6). The center of the barrel forms a pore lined by α-helices or β-strands. The side chains of the amino acid residues comprising the α-helices or β-strands establish the charge (cation or anion) selectivity of the channel. The degree of selectivity, or what specific ions are allowed to pass through the channel, depends on the diameter of the narrowest part of the pore.

Ion Transporters

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient. These transmembrane ATPases are divided into three families. The vacuolar (V) class of ion transporters includes H^+ pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of H^+ pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and

inorganic phosphate (P_i). The phosphorylated (P) class ion transporters, including Na^+-K^+ ATPase, Ca^{+2}-ATPase, and H^+-ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of Na^+ and Ca^{+2} are low and cytosolic concentration of K^+ is high. The resting potential of the cell is utilized in many processes involving carrier proteins and gated ion channels. Carrier proteins utilize the resting potential to transport molecules into and out of the cell. Amino acid and glucose transport into many cells is linked to sodium ion co-transport (symport) so that the movement of Na^+ down an electrochemical gradient drives transport of the other molecule up a concentration gradient. Similarly, cardiac muscle links transfer of Ca^{+2} out of the cell with transport of Na^+ into the cell (antiport).

Gated Ion Channels

Gated ion channels control ion flow by regulating the opening and closing of pores. The ability to control ion flux through various gating mechanisms allows ion channels to mediate such diverse signaling and homeostatic functions as neuronal and endocrine signaling, muscle contraction, fertilization, and regulation of ion and pH balance. Gated ion channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open their pores in response to mechanical stress; voltage-gated channels (e.g., Na^+, K^+, Ca^{+2}, and Cl^- channels) open their pores in response to changes in membrane potential; and ligand-gated channels (e.g., acetylcholine-, serotonin-, and glutamate- gated cation channels, and GABA- and glycine- gated chloride channels) open their pores in the presence of a specific ion, nucleotide, or neurotransmitter. The gating properties of a particular ion channel (i.e., its threshold for and duration of opening and closing) are sometimes modulated by association with auxiliary channel proteins and/or post translational modifications, such as phosphorylation. The pore forming subunits of voltage-gated and transmitter-gated cation channels form two distinct superfamilies of conserved multipass membrane proteins.

Voltage-gated Na^+ and K^+ channels are necessary for the function of electrically excitable cells such as nerve, endocrine, and muscle cells. Action potentials, which lead to neurotransmitter release and muscle contraction, arise from large, transient changes in the permeability of the membrane to Na^+ and K^+ ions. Depolarization of the membrane beyond the threshold level opens voltage-gated Na^+ channels. Sodium ions flow into the cell, further depolarizing the membrane and opening more voltage-gated Na^+ channels, thus propagating the depolarization down the length of the cell. Depolarization also opens voltage-gated K^+ channels. Consequently, potassium ions flow outward, leading to repolarization of the membrane. Voltage-gated channels utilize charged residues in the fourth transmembrane segment (S4) to sense voltage change. The open

state lasts only about 1 millisecond, at which time the channel spontaneously converts into an inactive state that cannot be opened irrespective of the membrane potential. Inactivation is mediated by the channel's N-terminus, which acts as a plug that closes the pore. The transition from an inactive to a closed state requires a return to resting potential.

Na^+ channels isolated from rat brain tissue are heterotrimeric complexes composed of a 260 kDa pore-forming α subunit that associates with two smaller auxiliary subunits, $\beta1$ and $\beta2$. The $\beta2$ subunit is an integral membrane glycoprotein that contains an extracellular Ig domain, and its association with α and $\beta1$ subunits correlates with increased function of the channel, a change in the channel's gating properties, as well as an increase in whole cell capacitance (Isom, L.L. et al. (1995) Cell 83:433-442).

K^+ channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as Ca^{++} and cAMP. In non-excitable tissue, K^+ channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, K^+ channels are responsible for setting resting membrane potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell contains a Na^+-K^+ pump and ion channels that provide the redistribution of Na^+, K^+, and Cl^-. The pump actively transports Na^+ out of the cell and K^+ into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow K^+ and Cl^- to flow by passive diffusion. Because of the high negative charge within the cytosol, Cl^- flows out of the cell. The flow of K^+ is balanced by an electromotive force pulling K^+ into the cell, and a K^+ concentration gradient pushing K^+ out of the cell. Thus, the resting membrane potential is primarily regulated by K^+ flow (Salkoff, L. and T. Jegla (1995) Neuron 15:489-492).

K^+ pore-forming subunits generally have six transmembrane-spanning domains with a short region between the fifth and sixth transmembrane regions that senses membrane potential; and the amino and carboxy termini are located intracellularly. In mammalian heart, the duration of ventricular action potential is controlled by a K^+ current. Thus, the K^+ channel is central to the control of heart rate and rhythm. K^+ channel dysfunctions are associated with a number of renal diseases including hypertension, hypokalemia, and the associated Bartter's syndrome and Getelman's syndrome, as well as neurological disorders including epilepsy. K^+ channels have been implicated in Alzheimer's disease by observations that a significant component of senile plaques, beta amyloid or A beta, also blocks voltage-gated potassium channels in hippocampal neurons. (See Antes, L.M. et al. (1998) Seminar Nephrol. 18:31-45; Stoffel, M. and L.Y. Jan (1998) Nat. Genet. 18:6-8; Madeja, M. et al. (1997) Eur. J. Neurosci. 9:390-395; Good, T.A. et al.

(1996) Biophys. J. 70:296-304.)

Voltage-gated Ca^{+2} channels are involved in presynaptic neurotransmitter release, and heart and skeletal muscle contraction. The voltage-gated Ca^{+2} channels from skeletal muscle (L-type) and brain (N-type) have been purified and, though their functions differ dramatically, they have similar subunit compositions. The channels are composed of three subunits. The α_1 subunit forms the membrane pore and voltage sensor, while the $\alpha_2\delta$ and β subunits modulate the voltage-dependence, gating properties, and the current amplitude of the channel. These subunits are encoded by at least six α_1, one $\alpha_2\delta$, and four β genes. A fourth subunit, γ, has been identified in skeletal muscle. (See Walker, D. et al. (1998) J. Biol. Chem. 273:2361-2367; Jay, S.D. et al. (1990) Science 248:490-492.)

Chloride channels are necessary in endocrine secretion and in regulation of cytosolic and organelle pH. In secretory epithelial cells, Cl^- enters the cell across a basolateral membrane through an Na^+, K^+/Cl^- cotransporter, accumulating in the cell above its electrochemical equilibrium concentration. Secretion of Cl^- from the apical surface, in response to hormonal stimulation, leads to flow of Na^+ and water into the secretory lumen. The cystic fibrosis transmembrane conductance regulator (CFTR) is a chloride channel encoded by the gene for cystic fibrosis, a common fatal genetic disorder in humans. Loss of CFTR function decreases transepithelial water secretion and, as a result, the layers of mucus that coat the respiratory tree, pancreatic ducts, and intestine are dehydrated and difficult to clear. The resulting blockage of these sites leads to pancreatic insufficiency, "meconium ileus," and devastating "chronic obstructive pulmonary disease" (Al-Awqati, Q. et al. (1992) J. Exp. Biol. 172:245-266).

Many intracellular organelles contain H^+-ATPase pumps that generate transmembrane pH and electrochemical differences by moving protons from the cytosol to the organelle lumen. If the membrane of the organelle is permeable to other ions, then the electrochemical gradient can be abrogated without affecting the pH differential. In fact, removal of the electrochemical barrier allows more H^+ to be pumped across the membrane, increasing the pH differential. Cl^- is the sole counterion of H^+ translocation in a number of organelles, including chromaffin granules, Golgi vesicles, lysosomes, and endosomes. Functions that require a low vacuolar pH include uptake of small molecules such as biogenic amines in chromaffin granules, processing of vacuolar constituents such as pro-hormones by proteolytic enzymes, and protein degradation in lysosomes (Al-Awqati, supra).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic

membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of Na^+ and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as γ-aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential.

Ligand-gated channels can be regulated by intracellular second messengers. Calcium-activated K^+ channels are gated by internal calcium ions. In nerve cells, an influx of calcium during depolarization opens K^+ channels to modulate the magnitude of the action potential (Ishi, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656). Cyclic nucleotide-gated (CNG) channels are gated by cytosolic cyclic nucleotides. The best examples of these are the cAMP-gated Na^+ channels involved in olfaction and the cGMP-gated cation channels involved in vision. Both systems involve ligand-activation of a G-protein coupled receptor which then alters the level of cyclic nucleotide within the cell. In olfaction, binding of an odorant to the receptor activates adenylate cyclase, leading to a rise in cytosolic cAMP. The cAMP binds to the cAMP-gated Na^+ channel causing an influx of Na^+, depolarization of the membrane, and initiation of a nerve impulse that travels along the axon to the brain. In vision, light activation of rhodopsin leads to activation of cGMP phosphodiesterase, which hydrolyzes cGMP. As a result, cytosolic cGMP levels drop, cGMP dissociates from cGMP-gated cation channels, and the channels close, resulting in hyperpolarization of the membrane. (See Zagotta, W.M. and S.A. Siegelbaum (1996) Annu. Rev. Neurosci. 19:235-263; Molday, R.S. and L.L. Molday (1998) Vision Res. 38:1315-1323.)

The subunits or monomers of an ion channel may be identical or different. CNG channels, for example, consist of α and β subunits that differ from each other at the N-terminal cytoplasmic tail. The central pore formed by the barrel arrangement is lined by an antiparallel β-sheet, the pore (P) region, contained within each subunit. This region also contains information specifying the ion selectivity for the channel. In the case of K^+ channels, a GYG tripeptide is involved in this selectivity (Ishi et al., supra). In voltage-gated channels, one of the transmembrane domains contains regularly spaced, positively charged amino acids that act as a voltage-sensor. In CNG channels, a region in the C-terminal cytoplasmic domain acts as a cyclic nucleotide binding site (Zagotta and Siegelbaum, supra). Ion channels also have a domain that functions in inactivation of the channel. In CNG K^+ channels, the inactivation domain is on the N-terminal cytoplasmic tail of the β-subunit. This domain acts as a tethered ball to block ion flow through the pore. This domain is also expressed as a separate protein, a glutamic acid-rich protein

(GARP), by alternative splicing and may act as an independent regulator of pore activity (Sautter, A. et al. (1997) Molec. Brain Res. 48:171-175).

Ion channels are essential to a wide range of physiological functions including neuronal signaling, muscle contraction, cardiac pacemaking, hormone secretion, and cell proliferation. Ion channels are expressed in a number of tissues where they are implicated in a variety of processes. CNG channels, while abundantly expressed in photoreceptor and olfactory sensory cells, are also found in kidney, lung, pineal, retinal ganglion cells, testis, aorta, and brain. Calcium-activated K^+ channels may be responsible for the vasodilatory effects of bradykinin in the kidney and for shunting excess K^+ from brain capillary endothelial cells into the blood. They are also implicated in repolarizing granulocytes after agonist-stimulated depolarization (Ishi et al., supra). Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol. 39:47-98).

The discovery of new human membrane channel proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative, immune/inflammatory, transport/secretory, osmoregulatory, muscular, cardiovascular, and neurological disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human membrane channel proteins, referred to collectively as "MECHP" and individually as "MECHP-1," "MECHP-2," "MECHP-3," "MECHP-4," "MECHP-5," "MECHP-6," "MECHP-7," "MECHP-8," "MECHP-9," "MECHP-10," "MECHP-11," "MECHP-12," "MECHP-13," "MECHP-14," "MECHP-15," "MECHP-16", "MECHP-17", and "MECHP-18." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18, and fragments thereof.

The invention further provides a substantially purified variant having at least 95% amino acid sequence identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-18 and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 95% polynucleotide sequence

identity to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36, and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 95% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36 and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36 and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected

from the group consisting of SEQ ID NO:1-18 and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of MECHP, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of MECHP, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-18 and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Figure 1 shows the amino acid sequence alignment between MECHP-1 (1568324; SEQ ID NO:1) and rat glutamic acid-rich protein (GI 2924369; SEQ ID NO:37), produced using the BLAST search tool.

Figure 2 shows the amino acid sequence alignment among MECHP-2 (4094907; SEQ ID NO:2), Drosophila voltage-gated potassium channel (GI 116443; SEQ ID NO:38), and P. penicillatus potassium channel α-subunit (GI 1763619; SEQ ID NO:39), produced using the multisequence alignment program of LASERGENE software (DNASTAR, Madison WI).

Figures 3A and 3B show the amino acid sequence alignment between MECHP-3 (518158; SEQ ID NO:3) and rat calcium-activated potassium channel rSK3 (GI 2564072; SEQ ID NO:40), produced using the multisequence alignment program of LASERGENE software.

Figures 4A, 4B, and 4C show the amino acid sequence alignment among MECHP-4 (602926; SEQ ID NO:4), Drosophila voltage-gated potassium channel (GI 116443; SEQ ID NO:38) and P. penicillatus potassium channel α-subunit (GI 1763619; SEQ ID NO:39), produced using the multisequence alignment program of LASERGENE software.

Figures 5A and 5B show the amino acid sequence alignment between MECHP-5 (922119; SEQ ID NO:5) and rat aquaporin 7 (GI 2350843; SEQ ID NO:41), produced using the multisequence alignment program of LASERGENE software.

Figures 6A and 6B show the amino acid sequence alignment between MECHP-7 (2731369; SEQ ID NO:7) and mouse connexin 30.3 (GI 192647; SEQ ID NO:42), produced using the multisequence alignment program of LASERGENE software.

Figure 7 shows the amino acid sequence alignment between MECHP-16 (2069907; SEQ ID NO:16) and human beta subunit of Ca^{+} activated K^{+} channel (GI 1055345; SEQ ID NO:43), produced using the multisequence alignment program of LASERGENE software.

Figures 8A and 8B show the amino acid sequence alignment between MECHP-17 (2243917; SEQ ID NO:17) and a homolog of Caenorhabditis elegans K^{+} channel protein (GI 3292929; SEQ ID NO:44), produced using the multisequence alignment program of LASERGENE software.

Figures 9A and 9B show the amino acid sequence alignment between MECHP-18 (2597476; SEQ ID NO:18) and human aquaporin 9 (GI 2887407; SEQ ID NO:45), produced using the multisequence alignment program of LASERGENE software.

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding MECHP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods and algorithms used for identification of MECHP.

Table 3 shows the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis, diseases, disorders, or conditions associated with these tissues, and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding MECHP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze MECHP, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"MECHP" refers to the amino acid sequences of substantially purified MECHP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to MECHP, increases or prolongs the duration of the effect of MECHP. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of MECHP.

An "allelic variant" is an alternative form of the gene encoding MECHP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding MECHP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as MECHP or a polypeptide with at least one functional characteristic of MECHP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding MECHP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding MECHP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent MECHP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity,

hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of MECHP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of MECHP which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of MECHP. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to MECHP, decreases the amount or the duration of the effect of the biological or immunological activity of MECHP. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the effect of MECHP.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind MECHP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on

the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic MECHP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding MECHP or fragments of MECHP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping

sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding MECHP, by northern analysis is indicative of the presence of nucleic acids encoding MECHP in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding MECHP.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR)

which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) Parameters for each method may be the default parameters provided by MEGALIGN or may be specified by the user. The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which

may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of MECHP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of MECHP.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:19-36, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:19-36 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:19-36 from related polynucleotide sequences. A fragment of SEQ ID NO:19-36 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:19-36 and the region of SEQ ID NO:19-36 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" and "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which

comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding MECHP, or fragments thereof, or MECHP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of

foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of MECHP polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to MECHP. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human membrane channel proteins (MECHP), the polynucleotides encoding MECHP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, immune/inflammatory, transport/secretory, osmoregulatory, muscular, cardiovascular, and neurological disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding

MECHP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each MECHP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The clones in column 5 were used to assemble the consensus nucleotide sequence of each MECHP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows the identity of each polypeptide; and column 7 shows analytical methods used to identify each polypeptide through sequence homology and protein motifs.

MECHP-1 has chemical and structural similarity with rat glutamic acid-rich protein (GI 2924369; SEQ ID NO:37). In particular, MECHP-1 and rat glutamic acid-rich protein share 15% overall identity. As shown in Figure 1, BLAST analysis identifies regions of MECHP-1 and rat glutamic acid-rich protein which share 27-30% identity. These regions extend from residue V12 through T163, P266 through G344, P461 through E548, and E653 through G709 in MECHP-1.

As shown in Figure 2, MECHP-2 has chemical and structural similarity with Drosophila voltage-gated potassium channel (GI 116443; SEQ ID NO:38) and P. penicillatus potassium channel α-subunit (GI 1763619; SEQ ID NO:39). In particular, MECHP-2 shares 18% identity with Drosophila voltage-gated K^+ channel, and 17% identity with P. penicillatus K^+ channel α-subunit. In particular, MECHP-2 shares 27% identity with Drosophila voltage-gated potassium channel and P. penicillatus potassium channel α-subunit over the first 133 residues, from M1 through T133 in MECHP-2.

As shown in Figures 3A and 3B, MECHP-3 has chemical and structural similarity with rat calcium-activated potassium channel rSK3 (GI 2564072; SEQ ID NO:40). In particular, MECHP-3 and rat rSK3 share 40% identity. MECHP-3 and rat rSK3 also share a canonical ion pore (P) region, including a GYG potassium ion selectivity sequence, from residue W192 through G213 in MECHP-3.

As shown in Figures 4A, 4B, and 4C, MECHP-4 has chemical and structural similarity with Drosophila voltage-gated potassium channel (GI 116443; SEQ ID NO:38) and P. penicillatus potassium channel α-subunit (GI 1763619; SEQ ID NO:39). In particular, MECHP-4 shares 28%

identity with Drosophila voltage-gated K^+ channel, and 26% identity with P. penicillatus K^+ channel α-subunit, respectively. MECHP-4, Drosophila voltage-gated K^+ channel, and P. penicillatus K^+ channel α-subunit also share a GYG potassium ion selectivity sequence from residue G372 through G374 in MECHP-4.

As shown in Figures 5A and 5B, MECHP-5 has chemical and structural similarity with rat aquaporin 7 (GI 2350843; SEQ ID NO:41). In particular, MECHP-5 and rat aquaporin 7 share 74% identity.

As shown in Figures 6A and 6B, MECHP-7 has chemical and structural similarity with mouse connexin 30.3 (GI 192647; SEQ ID NO:42). In particular, MECHP-7 and mouse connexin 30.3 (GI 192647) share 84% identity.

As shown in Figure 7, MECHP-16 has chemical and structural similarity with human beta subunit of Ca^+ activated K^+ channel (GI 1055345; SEQ ID NO:43). In particular, MECHP-16 and human beta subunit of Ca^+ activated K^+ channel share 40% identity.

As shown in Figures 8A and 8B, MECHP-17 has chemical and structural similarity with a homolog of C. elegans K^+ channel protein (GI 3292929; SEQ ID NO:44). In particular, MECHP-17 and the specified homolog of C. elegans K^+ channel protein share 47% identity.

As shown in Figures 9A and 9B, MECHP-18 has chemical and structural similarity with human aquaporin 9 (GI 2887407; SEQ ID NO:45). In particular, MECHP-18 and human aquaporin 9 share 46% identity.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding MECHP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists tissue categories which express MECHP as a fraction of total tissue categories expressing MECHP. Column 3 lists diseases, disorders, or conditions associated with those tissues expressing MECHP. Column 4 lists the vectors used to subclone the cDNA library. Northern analysis shows the expression of SEQ ID NO:34 in only 7 libraries, of which 6 (86%) are associated with cell proliferation. Two of these libraries are associated with brain tissue, one with pancreatic islet cells, one with kidney tissue, one with fetal lung tissue, one with ovarian tissue, and one with adrenal tissue. Northern analysis shows the expression of SEQ ID NO:36 in only 3 libraries, one of which is associated with ovarian tumor tissue, one with developing lung tissue, and one with gastrointestinal tissue associated with inflammation. Of particular note is the enriched expression of MECHP in neural and neuroendocrine tissue, most prominently the neural tissue-specific expression of SEQ ID NO:30.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding MECHP were isolated. Column 1 references the

nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The following fragments of the nucleotide sequences encoding MECHP are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:19-36, and to distinguish between SEQ ID NO:19-36 and related polynucleotide sequences. The useful fragments include the fragment of SEQ ID NO:19 from about nucleotide 764 to about nucleotide 808; the fragment of SEQ ID NO:20 from about nucleotide 523 to about nucleotide 582; the fragment of SEQ ID NO:21 from about nucleotide 628 to about nucleotide 669; the fragment of SEQ ID NO:22 from about nucleotide 779 to about nucleotide 826; the fragment of SEQ ID NO:23 from about nucleotide 64 to about nucleotide 108; the fragment of SEQ ID NO:24 from about nucleotide 1133 to about nucleotide 1180; the fragment of SEQ ID NO:25 from about nucleotide 656 to about nucleotide 700; the fragment of SEQ ID NO:26 from about nucleotide 153 to about nucleotide 197; the fragment of SEQ ID NO:27 from about nucleotide 2160 to about nucleotide 2219; the fragment of SEQ ID NO:28 from about nucleotide 1275 to about nucleotide 1322; the fragment of SEQ ID NO:29 from about nucleotide 313 to about nucleotide 348; the fragment of SEQ ID NO:30 from about nucleotide 994 to about nucleotide 1041; the fragment of SEQ ID NO:31 from about nucleotide 443 to about nucleotide 478; the fragment of SEQ ID NO:32 from about nucleotide 1175 to about nucleotide 1207; the fragment of SEQ ID NO:34 from about nucleotide 381 to about nucleotide 425; the fragment of SEQ ID NO:35 from about nucleotide 17 to about nucleotide 61; and the fragment of SEQ ID NO:36 from about nucleotide 54 to about nucleotide 98. The polypeptides encoded by the fragments of SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:34, SEQ ID NO:35, AND SEQ ID NO:36 are useful, for example, as immunogenic peptides.

The invention also encompasses MECHP variants. A preferred MECHP variant is one which has at least about 80%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the MECHP amino acid sequence, and which contains at least one functional or structural characteristic of MECHP.

The invention also encompasses polynucleotides which encode MECHP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:19-36, which encodes MECHP.

The invention also encompasses a variant of a polynucleotide sequence encoding MECHP. In particular, such a variant polynucleotide sequence will have at least about 70%, more

preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding MECHP. A particular aspect of the invention encompasses a variant of a sequence selected from the group consisting of SEQ ID NO:19-36 which has at least about 70%, more preferably at least about 85%, and most preferably at least about 95% polynucleotide sequence identity to a sequence selected from the group consisting of SEQ ID NO:19-36. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of MECHP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding MECHP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring MECHP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode MECHP and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring MECHP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding MECHP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding MECHP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode MECHP and MECHP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding MECHP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:19-36, or to a fragment of SEQ ID NO:19-36, under various conditions of stringency. (See,

e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or

combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the Robbins Hydra microdispenser (Robbins Scientific, Sunnyvale CA), Hamilton MICROLAB 2200 (Hamilton, Reno NV), Peltier Thermal Cycler 200 (PTC200; MJ Research, Watertown MA) and the ABI CATALYST 800 (Perkin-Elmer). Sequencing is then carried out using either ABI 373 or 377 DNA sequencing systems (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding MECHP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include

sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode MECHP may be cloned in recombinant DNA molecules that direct expression of MECHP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express MECHP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter MECHP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding MECHP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223, and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, MECHP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of MECHP, or any part thereof, may be altered during direct synthesis and/or

combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active MECHP, the nucleotide sequences encoding MECHP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding MECHP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding MECHP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding MECHP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding MECHP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding MECHP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression

vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding MECHP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding MECHP can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or pSPORT1 plasmid (Life Technologies). Ligation of sequences encoding MECHP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of MECHP are needed, e.g. for the production of antibodies, vectors which direct high level expression of MECHP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of MECHP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of MECHP. Transcription of sequences encoding MECHP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding MECHP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses MECHP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of MECHP in cell lines is preferred. For example, sequences encoding MECHP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* or *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* or *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its

substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding MECHP is inserted within a marker gene sequence, transformed cells containing sequences encoding MECHP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding MECHP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding MECHP and that express MECHP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of MECHP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on MECHP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding MECHP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding MECHP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an

appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding MECHP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode MECHP may be designed to contain signal sequences which direct secretion of MECHP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38), are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding MECHP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric MECHP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of MECHP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A

fusion protein may also be engineered to contain a proteolytic cleavage site located between the MECHP encoding sequence and the heterologous protein sequence, so that MECHP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled MECHP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of MECHP may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin-Elmer). Various fragments of MECHP may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of MECHP and human membrane channel proteins. In addition, the expression of MECHP is closely associated with nervous, reproductive, and gastrointestinal tissues; fetal development; and neurological, immune/inflammatory, and cell proliferative disorders, including cancer. Therefore, MECHP appears to play a role in cell proliferative, immune/inflammatory, transport/secretory, osmoregulatory, muscular, cardiovascular, and neurological disorders. In the treatment of disorders associated with increased MECHP expression or activity, it is desirable to decrease the expression or activity of MECHP. In the treatment of disorders associated with decreased MECHP expression or activity, it is desirable to increase the expression or activity of MECHP.

Therefore, in one embodiment, MECHP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MECHP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma,

leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a transport/secretory disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, Chediak-Higashi syndrome, diabetes mellitus, diabetes insipidus, diabetic neuropathy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, malignant hyperthermia, multidrug resistance, myotonic dystrophy, catatonia, dystonias, peripheral neuropathy, neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, toxic shock syndrome, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, goiter, Cushing's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, and allergies, including hay fever, asthma, and urticaria (hives); an osmoregulatory disorder such as diabetes insipidus, diarrhea, peritonitis, chronic renal failure, Addison's disease, SIADH, hypoaldosteronism, hyponatremia, adrenal insufficiency, hypothyroidism, hypernatremia, hypokalemia, Barter's syndrome, metabolic acidosis, metabolic alkalosis, encephalopathy, edema, hypotension, and hypertension; a muscular disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy,

and ethanol myopathy; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation; congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Down syndrome, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases; neuromuscular disorders including spinal

muscular atrophy, carpal tunnel syndrome, monomeuritis multiplex; muscular dystrophies such as Duchenne's, myotonic facioscapulohumeral, oculopharyngeal, scapuloperoneal, congenital, distal, and ocular; congenital and metabolic myopathies, myotonia, peripheral nervous system disorders, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including depression and bipolar disorder, and mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; abnormalities in electrolytes such as calcium, phosphate, magnesium, and potasium; hypo- and hyperfunction of the thyroid, adrenal, parathyroid, and pituitary; and primary and metastatic neoplasms.

In another embodiment, a vector capable of expressing MECHP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MECHP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified MECHP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MECHP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of MECHP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of MECHP including, but not limited to, those listed above.

In a further embodiment, an antagonist of MECHP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MECHP. Such disorders may include, but are not limited to, those discussed above. In one aspect, an antibody which specifically binds MECHP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express MECHP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding MECHP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MECHP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment

or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of MECHP may be produced using methods which are generally known in the art. In particular, purified MECHP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind MECHP. Antibodies to MECHP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with MECHP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to MECHP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of MECHP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to MECHP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984)

Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce MECHP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for MECHP may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between MECHP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering MECHP epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for MECHP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of MECHP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple MECHP epitopes, represents the average affinity, or avidity, of the antibodies for MECHP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular MECHP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} l/mole are preferred for use in immunoassays in which the MECHP-antibody complex must withstand

rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 l/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of MECHP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of MECHP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding MECHP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding MECHP may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding MECHP. Thus, complementary molecules or fragments may be used to modulate MECHP activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding MECHP.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding MECHP. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding MECHP can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding MECHP. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing

complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding MECHP. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding MECHP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding MECHP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs

and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotech. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of MECHP, antibodies to MECHP, and mimetics, agonists, antagonists, or inhibitors of MECHP. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets,

pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of MECHP, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example MECHP or fragments thereof, antibodies of MECHP, and agonists, antagonists or inhibitors of MECHP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind MECHP may be used for the diagnosis of disorders characterized by expression of MECHP, or in assays to monitor patients being treated with MECHP or agonists, antagonists, or inhibitors of MECHP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for MECHP include methods which utilize the antibody and a label to detect MECHP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring MECHP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of MECHP expression. Normal or standard values for MECHP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to MECHP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of MECHP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding MECHP may be

used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of MECHP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of MECHP, and to monitor regulation of MECHP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding MECHP or closely related molecules may be used to identify nucleic acid sequences which encode MECHP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding MECHP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the MECHP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:19-36 or from genomic sequences including promoters, enhancers, and introns of the gene encoding MECHP.

Means for producing specific hybridization probes for DNAs encoding MECHP include the cloning of polynucleotide sequences encoding MECHP or MECHP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding MECHP may be used for the diagnosis of disorders associated with expression of MECHP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas,

parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an immune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a transport/secretory disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, Chediak-Higashi syndrome, diabetes mellitus, diabetes insipidus, diabetic neuropathy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, malignant hyperthermia, multidrug resistance, myotonic dystrophy, catatonia, dystonias, peripheral neuropathy, neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, toxic shock syndrome, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, goiter, Cushing's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, and allergies, including hay fever, asthma, and urticaria (hives); an osmoregulatory disorder such as diabetes insipidus, diarrhea, peritonitis, chronic renal failure, Addison's disease, SIADH, hypoaldosteronism, hyponatremia, adrenal insufficiency, hypothyroidism, hypernatremia, hypokalemia, Barter's syndrome, metabolic acidosis, metabolic alkalosis, encephalopathy, edema, hypotension, and hypertension; a muscular disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, and ethanol myopathy; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis,

balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation; congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Down syndrome, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases; neuromuscular disorders including spinal muscular atrophy, carpal tunnel syndrome, monomeuritis multiplex; muscular dystrophies such as Duchenne's, myotonic facioscapulohumeral, oculopharyngeal, scapuloperoneal, congenital, distal, and ocular; congenital and metabolic myopathies, myotonia, peripheral nervous system disorders,

dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis; mental disorders including depression and bipolar disorder, and mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; abnormalities in electrolytes such as calcium, phosphate, magnesium, and potasium; hypo- and hyperfunction of the thyroid, adrenal, parathyroid, and pituitary; and primary and metastatic neoplasms. The polynucleotide sequences encoding MECHP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered MECHP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding MECHP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding MECHP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding MECHP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of MECHP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding MECHP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period

ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or over-expressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding MECHP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding MECHP, or a fragment of a polynucleotide complementary to the polynucleotide encoding MECHP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of MECHP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding MECHP may be used to generate hybridization probes useful in mapping the naturally occurring genomic

sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding MECHP on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, MECHP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between MECHP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test

compounds are synthesized on a solid substrate. The test compounds are reacted with MECHP, or fragments thereof, and washed. Bound MECHP is then detected by methods well known in the art. Purified MECHP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding MECHP specifically compete with a test compound for binding MECHP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with MECHP.

In additional embodiments, the nucleotide sequences which encode MECHP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0589 P, filed September 2, 1998], U.S. Ser. No. [Attorney Docket No. PF-0632 P, filed November 12, 1998], U.S. Ser. No. [Attorney Docket No. PF-0648 P, filed December 9, 1998], U.S. Ser. No. [Attorney Docket No. PF-0664 P, filed January 26, 1999], and U.S. Ser. No. [Attorney Docket No. PF-0671 P, filed February 10, 1999], are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA

purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), pSPORT1 plasmid (Life Technologies), or pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids were recovered from host cells by in vivo excision, using the UNIZAP vector system (Stratagene) or cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a Fluoroskan II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing systems (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open

reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:19-36. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals, Palo Alto CA). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\ \text{sequence identity} \times \%\ \text{maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding MECHP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer,

inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in the description of the invention.

V. Extension of MECHP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:19-36 was produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICO GREEN quantitation reagent (0.25% (v/v) PICO GREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates,

digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with *Pfu* DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequence of SEQ ID NO:19-36 is used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:19-36 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [^{32}P]-adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba1, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon

membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the MECHP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring MECHP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of MECHP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the MECHP-encoding transcript.

IX. Expression of MECHP

Expression and purification of MECHP are achieved using bacterial or virus-based expression systems. For expression of MECHP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express MECHP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MECHP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding MECHP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E. K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, MECHP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from MECHP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified MECHP obtained by these methods can be used directly in the following activity assay.

X. Demonstration of MECHP Activity

Aquaporin Activity of MECHP

Aquaporin activity of MECHP is demonstrated as the ability to induce osmotic water permeability in Xenopus laevis oocytes injected with MECHP cRNA (Ishibashi, K. et al. (1994)

Proc. Natl. Acad. Sci. USA 91:6269-6273). Oocytes injected with water are used as the control. Injected oocytes are given a hypotonic shock by being transferred from 200 mosM to 70 mosM modified Barth's buffer. The increase in osmotic volume of the oocytes, observed at 24°C by videomicroscopy, is proportional to the MECHP aquaporin activity in the injected oocytes.

Protein Transport Activity of MECHP

Protein transport activity of MECHP is demonstrated by its ability to catalyze the translocation of newly synthesized preprolactin into proteoliposomes in an in vitro system (Görlich, D. and T.A. Rapoport (1993) Cell 75:615-630). Proteoliposomes are prepared containing purified MECHP, purified dog Sec61p beta and gamma, purified dog SRP receptor, and a mixture of phospholipids (phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, and phosphatidylinositol) corresponding approximately to those found in native microsomes. The proteoliposomes are incubated in a wheat germ in vitro translation system in which a secretory protein (preprolactin) is synthesized in the presence of SRP and radioactive amino acids. After translation and synthesis of preprolactin, half of the sample is treated with 500 μg/ml proteinase K while the other half remains untreated. Any translocated preprolactin will be inaccessible to proteinase K while any untranslocated preprolactin will be degraded. The amount of preprolactin in the samples with and without proteinase K treatment is determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis followed by phosphor image analysis. The amount of preprolactin protected from proteinase K digestion in the proteinase K-treated sample is proportional to the protein transport activity of MECHP.

Gap Junction Activity of MECHP

Gap junction activity of MECHP is demonstrated as the ability to induce the formation of intercellular channels between paired Xenopus laevis oocytes injected with MECHP cRNA (Hennemann, supra). One week prior to the experimental injection with MECHP cRNA, oocytes are injected with antisense oligonucleotide to MECHP to reduce background. MECHP cRNA-injected oocytes are incubated overnight, stripped of vitelline membranes, and paired for recording of junctional currents by dual cell voltage clamp. The measured conductances are proportional to gap junction activity of MECHP.

Ion Channel Activity of MECHP

Ion channel activity of MECHP is demonstrated using an electrophysiological assay for ion conductance. MECHP can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding MECHP. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A second plasmid which expresses any one of a number of

marker genes, such as ß-galactosidase, is co-transformed into the cells to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of MECHP and ß-galactosidase.

Transformed cells expressing ß-galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells are tested for differences in membrane conductance due to potassium ions by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or ß-galactosidase sequences alone, are used as controls and tested in parallel. Cells expressing MECHP will have higher cation conductance relative to control cells. The contribution of MECHP to conductance can be confirmed by incubating the cells using antibodies specific for MECHP. The antibodies will bind to the extracellular side of MECHP, thereby blocking the pore in the ion channel, and the associated conductance.

Ion channel activity of MECHP is also measured as current flow across a MECHP-containing Xenopus oocyte membrane using the two-electrode voltage-clamp technique (Ishi et al., supra; Jegla, T. and L. Salkoff (1997) J. Neurosci. 17:32-44). MECHP is subcloned into an appropriate Xenopus oocyte expression vector, such as pBF, and 0.5-5 ng of mRNA is injected into mature stage IV oocytes. Injected oocytes are incubated at 18°C for 1-5 days. Inside-out macropatches are excised into an intracellular solution containing 116 mM K-gluconate, 4 mM KCl, and 10 mM Hepes (pH 7.2). The intracellular solution is supplemented with varying concentrations of the MECHP mediator, such as cAMP, cGMP, or Ca^{+2} (in the form of $CaCl_2$), where appropriate. Electrode resistance is set at 2-5 MΩ and electrodes are filled with the intracellular solution lacking mediator. Experiments are performed at room temperature from a holding potential of 0 mV. Voltage ramps (2.5 s) from -100 to 100 mV are acquired at a sampling frequency of 500 Hz. Current measured is proportional to the activity of MECHP in the assay.

XI. Functional Assays

MECHP function is assessed by expressing the sequences encoding MECHP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or

hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP, and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of MECHP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding MECHP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding MECHP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of MECHP Specific Antibodies

MECHP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the MECHP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A peptide

synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring MECHP Using Specific Antibodies

Naturally occurring or recombinant MECHP is substantially purified by immunoaffinity chromatography using antibodies specific for MECHP. An immunoaffinity column is constructed by covalently coupling anti-MECHP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing MECHP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MECHP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MECHP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and MECHP is collected.

XIV. Identification of Molecules Which Interact with MECHP

MECHP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MECHP, washed, and any wells with labeled MECHP complex are assayed. Data obtained using different concentrations of MECHP are used to calculate values for the number, affinity, and association of MECHP with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	19	1568324	UTRSNOT05	668081H1 (SCORNOT01), 777659H1 (COLNNOT05), 1419526F1 (KIDNNOT09), 1568324F6 (UTRSNOT05), 1568324H1 (UTRSNOT05), 1691082F6 (PROSTUT10), 1866748F6 (SKINBIT01), 2133242H1 (ENDCNOT01), 2740849H1 (BRSTTUT14), 3488431H1 (EPIGNOT01), 3534835H1 (KIDNNOT25), 3556331H1 (LUNGNOT31), 3747086H1 (THYMNOT08)
2	20	4094907	BSCNSZT01	1298228F6 (BRSTNOT07), 1298228T6 (BRSTNOT07), 3518650T6 (LUNGNON03), 1673339T6 (BLADNOT05), 4094907H1 (BSCNSZT01), 810976R1 (LUNGNOT04)
3	21	518158	MMLR1DT01	518158H1 (MMLR1DT01), 1322305X302F1 (BLADNOT04), 1339742F1 (COLNTUT03), 1662883F6 (BRSTNOT09), 1868856F6 (SKINBIT01), 3329796H1 (HEAONOT04), SAPA00287F1
4	22	602926	BRSTTUT01	602926H1 (BRSTTUT01), 602926R1 (BRSTTUT01), 602926X15 (BRSTTUT01), 602926X18 (BRSTTUT01), 1236735H1 (LUNGFET03), 1294713F6 (PGANNOT03), 1342719X29R1 (COLNTUT03), 1796484T6 (PROSTUT05)
5	23	922119	RATRNOT02	922119H1 (RATRNOT02), 2304391T6 (BRSTNOT05), 2925760H1 (SININOT04), 2925760T6 (SININOT04), 3283088H1 (HEAONOT05), 5330728H1 (DRGTNON04), 5343411H1 (CONFNOT05)
6	24	2666782	THYMFET03	2666782H1 (THYMFET03), 2666782X305D2 (THYMFET03), 2666782X310F2 (THYMFET03), 2998445H1 (OVARTUT07), 2999052F6 (OVARTUT07), 3244028F6 (BRAINOT19), 3244028X317B2 (BRAINOT19)
7	25	2731369	OVARTUT04	2631755F6 (COLNTUT15), 2631755X300D1 (COLNTUT15), 2631755X303B1 (COLNTUT15), 2631755X303D1 (COLNTUT15), 2731369H1 (OVARTUT04), 2798719F6 (NPOLNOT01), 4406377H1 (PROSDIT01)

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
8	26	1375415	LUNGNOT10	664378R6 (SCORNOT01), 1375415H1 (LUNGNOT10), 1477459H1 (CORPNOT02), 3068750H1 (UTRSNOR01)
9	27	2733282	OVARTUT04	1821059F6 (GBLATUT01), 2620171R6 (KERANOT02), 2620171X306U2 (KERANOT02), 2733282F6 (OVARTUT04), 2733282H1 (OVARTUT04), 2733282X315F1 (OVARTUT04), 3618886F6 (EPIPNOT01), SBLA01906F1, SBLA01292F1
10	28	3148427	ADRENON04	3148427H1 (ADRENON04), 3877333F6 (HEARNOT06), 3877333T6 (HEARNOT06), 3877333X331B1 (HEARNOT06), 3877333X331U1 (HEARNOT06)
11	29	3342358	SPLNNOT09	259592X14 (HNT2RAT01), 2481052H1 (SMCANOT01), 3342358H1 (SPLNNOT09)
12	30	1267774	BRAINOT09	1267774F6 (BRAINOT09), 1267774H1 (BRAINOT09), 1740673R6 (HIPONON01), 1740673T6 (HIPONON01), 3242923H1 (BRAINOT19), 4837609H1 (BRAWNOT01)
13	31	1817329	PROSNOT20	1817329F6 (PROSNOT20), 1817329H1 (PROSNOT20), 2506976F6 (CONUTUT01), 4313771F6 (BRAFNOT01), SAEA10065P1, SAEA02844F1, SAEA02136F1
14	32	3273307	PROSBPT06	2658420F6 (LUNGTUT09), 2658420X316D1 (LUNGTUT09), 2658420X325D1 (LUNGTUT09), 3273307F6 (PROSBPT06), 3273307H1 (PROSBPT06)
15	33	3824833	BRAXNOT01	3824833H1 (BRAXNOT01), SAGA02981F1, SAGA00581R1, SAGA01037F1
16	34	2069907	ISLTNOT01	2069907H1 (ISLTNOT01), 2069907X304D1 (ISLTNOT01), 2069907X313V1 (ISLTNOT01), 2736831F6 (OVARNOT09), 2736831T6 (OVARNOT09)
17	35	2243917	PANCTUT02	2243917H1 (PANCTUT02), 2243917F6 (PANCTUT02), 2108673R6 (BRAITUT03), 1804567F6 (SINTNOT13), 980106H1 (TONGTUT01)
18	36	2597476	OVARTUT02	2597476H1 (OVARTUT02), 2597476F6 (OVARTUT02), 1633918F6 (COLNNOT19), SAEC11415F1, SAEC10014F1, SBKA03380F1, SAEC10514F1

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods
1	724	S123 S212 S246 S8 S20 S26 S32 T36 S38 T48 S73 S81 S82 S102 S161 S215 S227 S309 S375 T400 S492 S511 S545 S551 S564 S586 S627 T642 S119 S136 S152 S206 S240 S278 S424 S439 S444 S528 S539 S597 S607 T608 Y249	N582	ATP/GTP-binding site motif A (P-loop): A460-S467	Glutamic acid-rich protein (cyclic nucleotide-gated cation channel subunit)	BLAST MOTIFS
2	257	S162 S70 T93 T133 T242		Potassium channel signature: H74-T93	Potassium channel	MOTIFS PRINTS
3	377	S284 S174 S317 T101 T279 S338	N182 N334	Canonical ion pore region: W192-G213 Potassium channel signatures: T184-V206, G213-L239 Signal peptide: M1-S68 Transmembrane domains: W28-M46, I65-A81, L154-L173	Calcium-activated potassium channel	BLAST HMM MOTIFS PRINTS SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods
4	491	S61 S144 S145 S150 S206 T274 S320 S463 S471 S472 T483 S75 T85 S118 T161 S183 T316 S489 Y437	N481 N487	Ion transport protein domain: C181-I405 Potassium channel signatures: E66-T85, P178-S206, G224-Q247, F250-L270, L294-S320, E323-E346, L354-T376, G383-F409 Potassium ion selectivity sequence: G372-G374 Transmembrane domain: V324-Y343	Voltage-gated potassium channel	BLAST HMM MOTIFS PFAM PRINTS
5	341	T189 S247 S9 Y292	N322	Major Intrinsic Protein (MIP): E26-Y271 Aquaporin NPA boxes: N93-A95, N225-S227 Transmembrane domain: M41-L59	Aquaporin	BLAST BLOCKS HMM MOTIFS PFAM PRINTS
6	476	T75 T105 T207 T222 S346 T378 S386 S71 T203 T224 S269 S309 Y235		Eubacterial secY protein: T75-I460 Signal peptide: M1-C46 Transmembrane domains: L33-F51, I147-L165, L239-I256	Sec61p alpha subunit	BLAST BLOCKS HMM MOTIFS PFAM PRINTS SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods
7	266	S115 S121 T181	N119 N201	Connexin: M1-L208 Signal peptide: M1-A39 Transmembrane domain: I23-A39	Connexin	BLAST BLOCKS HMM MOTIFS ProfileScan PFAM PRINTS SPScan
8	182	S9 S130	N108	Transmembrane domains: R36-I55, A65-F84, V103-S130	Voltage-gated K^+-channel (Plasmolipin)	BLAST HMM
9	942	S273 S302 S355 S368 T418 S419 S474 S498 T797 T2 T76 T92 S270 S318 S384 T557 T568 T593 T652 S685 T822 S823 S858 S927	N74 N97 N150 N231 N235 N253 N291 N521 N555 N579 N636 N821 N937	Signal peptide: M1-A31 Transmembrane domain: L900-L926	Calcium-dependent chloride channel (Lu-ECAM-1)	BLAST HMM SPScan
10	519	S74 S159 S187 T191 T224 T329 T441 S461 S466 S122 S172 T486 T516	N72 N215 N259 N394 N459	von Willebrand factor type A domain: D162-V337	L-type calcium channel subunit	BLAST PFAM PRINTS
11	251	S249 T52 T164 T182 S235 S35 S171	N50		Chloride intracellular channel (CLIC2)	BLAST

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods
12	323	S173 T321 S44 S46 S165 S290 T50 S51 S169 T211 S228 S240	N48	PMP-22/PM20/EMP family protein: V8-Q207, T59-N72, Y176-D203 Signal peptide: M1-G29 Transmembrane domains: M10-V28, I106-A123, I134-S158, F180-V198	Voltage-gated Ca^{2+} channel, gamma subunit	BLAST BLOCKS HMM MOTIFS PFAM SPScan
13	51	S5 T31			Ca^{2+} channel, beta subunit	BLAST MOTIFS
14	113	S101 S102 T12 T46 T108 T65 S95 S96			Ca^{2+}-activated K^{+} channel	BLAST MOTIFS
15	215	T68 S106 S192 T198 T204	N42 N66 N74	Immunoglobulin domain: G43-I129 Myelin PO protein signature: L92-P119, D121-E150, A159-V183 Signal peptide: M1-S29 Transmembrane domain: T157-L177	Na^{+} channel, beta subunit	BLAST HMM MOTIFS PFAM PRINTS SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences	Identification	Analytical Methods
16	235	T36 S90 S122 S176 S6 S12 T135	N88 N96 N119	Transmembrane domain: I48-T68	Ca^{2+}-activated K^{+} channel, beta subunit	BLAST HMM MOTIFS
17	234	S29 S82 T174 T216 T57 S221	N130	Microbodies C-terminal targeting signal: S232-M234	K^{+} channel protein	BLAST MOTIFS
18	301	T47 S286	N75 N128 N133	MIP family protein domain: R15-Y260 MIP family signature: H80-A88 Microbodies C-terminal targeting signal: C299-L301 Transmembrane domain: N54-Y71	Aquaporin 9	BLAST HMM MOTIFS PFAM

Table 3

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease, Disorder or Condition (Fraction of Total)	Vector
19	Reproductive Fetal	Cell Proliferative (0.570) Immune Response (0.210)	pINCY
20	Nervous Reproductive	Cell Proliferative (0.560) Immune Response (0.320)	pINCY
21	Reproductive Hematopoietic/Immune Gastrointestinal	Cell Proliferative (0.660) Immune Response (0.370)	PSPORT1
22	Reproductive Gastrointestinal Musculoskeletal	Cell Proliferative (0.790) Immune Response (0.210)	PSPORT1
23	Gastrointestinal (0.333) Reproductive (0.292) Cardiovascular (0.208)	Cell Proliferative and Cancer (0.625) Inflammation (0.500)	PSPORT1
24	Nervous (0.455) Reproductive (0.409)	Cell Proliferative and Cancer (0.591) Inflammation (0.273) Neurological (0.182)	pINCY
25	Nervous (0.400) Reproductive (0.400) Gastrointestinal (0.200)	Cell Proliferative and Cancer (0.800) Inflammation (0.200)	pINCY
26	Nervous (0.533) Gastrointestinal (0.133)	Cell Proliferative (0.533) Inflammation (0.267) Neurological (0.113)	pINCY
27	Reproductive (0.333) Dermatologic (0.167) Gastrointestinal (0.167)	Cell Proliferative (0.750)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Tissue Expression (Fraction of Total)	Disease, Disorder or Condition (Fraction of Total)	Vector
28	Cardiovascular (0.250) Endocrine (0.250) Nervous (0.250)	Cancer (0.429) Trauma (0.429)	PSPORT1
29	Cardiovascular (0.500) Gastrointestinal (0.250) Hematopoietic/Immune (0.250)	Cancer (0.500) Inflammation (0.250)	pINCY
30	Nervous (1.000)	Inflammation (0.334) Neurological (0.167)	pINCY
31	Reproductive (0.500) Nervous (0.300) Gastrointestinal (0.200)	Cancer (0.600) Neurological (0.100)	pINCY
32	Reproductive (0.286) Endocrine (0.286)	Cancer (0.500) Inflammation (0.500)	pINCY
33	Nervous (0.667) Reproductive (0.333)	Cancer (0.333) Inflammation (0.333)	pINCY
34	Nervous (0.286)	Cell Proliferative (0.857)	pINCY
35	Nervous (0.296) Gastrointestinal (0.259) Reproductive (0.185)	Cell Proliferative (0.593) Inflammation/Immune Response (0.222)	pINCY
36	Reproductive (0.333) Cardiovascular (0.333) Gastrointestinal (0.333)	Cell Proliferative and Cancer (0.677) Inflammation (0.333)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Comment
19	UTRSNOT05	Library was constructed using RNA isolated from the uterine tissue of a 45-year-old Caucasian female during a total abdominal hysterectomy and total colectomy. Pathology for the associated tumor tissue indicated multiple leiomyomas of the myometrium and a grade 2 colonic adenocarcinoma of the cecum. Patient history included multiple sclerosis and mitral valve disorder. Family history included type I diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.
20	BSCNSZT01	Library was constructed using RNA isolated from diseased caudate nucleus tissue removed from the brain of a 49-year-old male. Patient history included schizophrenia.
21	MMLR1DT01	Library was constructed using RNA isolated from adherent mononuclear cells, which came from a pool of male and female donors. The cells were cultured for 24 hours following Ficoll Hypaque centrifugation.
22	BRSTTUT01	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
23	RATRNOT02	Library was constructed using RNA isolated from the right atrium tissue of a 39-year-old Caucasian male, who died from a gunshot wound.
24	THYMFET03	Library was constructed using RNA isolated from thymus tissue removed from a Caucasian male fetus.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
25	OVARTUT04	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 53-year-old Caucasian female during a total abdominal hysterectomy, removal of the fallopian tubes and ovaries, regional lymph node excision, peritoneal tissue destruction, and incidental appendectomy. Pathology indicated grade 1 transitional cell carcinoma of the right ovary. The left ovary had a hemorrhagic corpus luteum. The uterus had multiple leiomyomas (1 submucosal, 11 intramural), and the endometrium was inactive. The cul-de-sac contained abundant histiocytes and rare clusters of mesothelial cells. Patient history included breast fibrosclerosis and chronic stomach ulcer. Family history included acute stomach ulcer with perforation, breast cancer, bladder cancer, rectal/anal cancer, benign hypertension, coronary angioplasty, and hyperlipidemia.
26	LUNGNOT10	Library was constructed using RNA isolated from lung tissue of a Caucasian male fetus, who died at 23 weeks' gestation.
27	OVARTUT04	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 53-year-old Caucasian female during a total abdominal hysterectomy, removal of the fallopian tubes and ovaries, regional lymph node excision, and peritoneal tissue destruction. Pathology indicated grade 1 transitional cell carcinoma of the right ovary. The left ovary had a hemorrhagic corpus luteum. The uterus had multiple leiomyomas (1 submucosal, 11 intramural), and the endometrium was inactive. The cul-de-sac contained abundant histiocytes and rare clusters of mesothelial cells. Patient history included breast fibrosclerosis and chronic stomach ulcer. Family history included acute stomach ulcer with perforation, breast cancer, bladder cancer, rectal/anal cancer, benign hypertension, coronary angioplasty, and hyperlipidemia.
28	ADRENON04	Library was constructed from 1.36 million independent clones from an adrenal tissue library. Starting RNA was made from adrenal gland tissue removed from a 20-year-old Caucasian male. The library was normalized in two rounds using conditions adapted from Soares et al. (Proc. Natl. Acad. Sci. USA (1994) 91:9228-9232) and Bonaldo et al. (Genome Res. (1996) 6:791-806), using a significantly longer (48-hours/round) reannealing hybridization period.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
29	SPLNNOT09	Library was constructed using RNA isolated from diseased spleen tissue removed from a 22-year-old Caucasian male (Ashkenazi Jewish descent) during a total splenectomy. Pathology indicated Gaucher's disease with marked splenomegaly. Patient history included thyroid disorders and type 1 Gaucher's disease. Family history included benign hypertension, thyroid disease, myocardial infarction, cerebrovascular disease, arteriosclerotic cardiovascular disease, and prostate cancer.
30	BRAINOT09	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
31	PROSNOT20	Library was constructed using RNA isolated from diseased prostate tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma.
32	PROSBPT06	Library was constructed using RNA isolated from diseased prostate tissue remove from a 66-year-old Caucasian male during a radical prostatectomy and lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated grade 2 (of 4) adenocarcinoma, Gleason grade 3+3. The patient presented with elevated prostate specific antigen (PSA), proteinuria, decreased renal function, and urinary frequency. Patient history included hemiparesis, depressive disorder, sleep apnea, psoriasis, mitral valve prolapse, cerebrovascular disease, benign hypertension, and impotence. Family history included benign hypertension, cerebrovascular disease, and colon cancer.
33	BRAXNOT01	Library was constructed using RNA isolated from cerebellar tissue removed from a 70-year-old male. Patient history included chronic obstructive airways disease and left ventricular failure.
34	ISLTNOT01	Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Comment
35	PANCTUT02	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreatic duodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.
36	OVARTUT02	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 51-year-old Caucasian female during an exploratory laparotomy, total abdominal hysterectomy, salpingo-oophorectomy, and an incidental appendectomy. Pathology indicated mucinous cystadenoma presenting as a multiloculated neoplasm involving the entire left ovary. The right ovary contained a follicular cyst and a hemorrhagic corpus luteum. The uterus showed proliferative endometrium and a single intramural leiomyoma. The peritoneal biopsy indicated benign glandular inclusions consistent with endosalpingiosis. Family history included atherosclerotic coronary artery disease, benign hypertension, breast cancer, and uterine cancer.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs*: Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs*: fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences*: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less where applicable
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score= 4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:16, SEQ ID NO:17, and SEQ ID NO:18, and fragments thereof.

2. A substantially purified variant having at least 95% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 95% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:19-36 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 95% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

(a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

(b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of MECHP, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of MECHP, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

```
12   VEPDPEAGSEQEVFSAVEGPSAEETPSDTESPEVLETQLDAHQGLLGMDPPGDMVDFVAA   1568324
     VPP      ++  A   PEP      PE EQ+        +P +    AA
8    VLPQPPGTPQKTEEGAGPQPETESKPEANPQPEP-EVQPEPEPEP-EPEPEPEPAPEEAA   GI 2924369

72   ESTEDL---KALSSEEEEEMGGAAQEPESLLPPSVLDQASVIAERFVSSFSRRSS-VAQE   1568324
       +L   ++ E+ EG+QE+  PP   QAV    V  +R SS+
66   PEVQTLPPEEPVEGEDVAEAGPSLQETQEADPPQPTSQAQVA----VVKVNRPSSWMLSW   GI 2924369

128  DSKSSGFGSPRLVSRSSSVLSLEGSEKGLARHGSAT                           1568324
      K      P+V SS  +L  EG +G+T
122  FWKGMEKVVPQPVYSSSGGQNLAAGEGGPDQDGAQT                           GI 2924369

461  PERDGKSPTVPCLQEEAGEPLGGKGKRK-----PVLSLFDYEQLMAQE--HSPPKPSSAG   1568324
     P++DG    PC +G  G    K     PLL + +L ++    PP PS A
150  PDQDGAQTLEPCGTGDPGSEDGSDKTSKTQDTEPSLWLLRWLELNLEKVLPQPPTPSQAW   GI 2924369

514  EMSPQRFFFNP-PAVSQRTTSPGGRPSARSPLSPTE                           1568324
     ++P+      PP +     P   PS +P PE
210  KVEPEGAVLEPDPPGTPMEVEPTENPSQPNP-GPVE                           GI 2924369

653  EKGPLPSPTAGLEESSGQGPSSPVALLGQVQDFQQSAECQPKEEGSRDPADPSQQG       1568324
     E+P  P G+SS  P PVL+ +  +A QP  G    +PS G
246  EEEPAAEPQPGFQASSLPPPGDPVRLIEWLLHRLEMALPQPVLHGKAAEQEPSCPG       GI 2924369
```

FIGURE 1

```
1   MSRPLITRSPAS--------PLNNQGIPTP 4094907
1   MAS-VAAWLPFARAAAIGWVPIATHPLPPP GI 116443
1   MNGDIGAWISCARTAGIGWVPISSKE-PSA GI 1763619

23  -------AQLTKSNAPVHIDVGGHMYTSSL 4094907
30  PMPKD---RRKTDDEKLLINVSGRRFETWR GI 116443
30  YLNKQVCNENEKNNAKLTINVSGRRYQTYS GI 1763619

46  ATLTKYPESRIGRLFDGTEPIVL-DSLKQH 4094907
57  NTLEKYPDT----LLGSNEREFFYDEDCKE GI 116443
60  HTLRKFKET----LLGSQERDYFYDESLEE GI 1763619

75  YFIDRDGQMFRYILNFLRTSKLLIPDDFKD 4094907
83  YFFDRDPDIFRHILNYYRTGKLHYPKH--E GI 116443
86  YYFDRDPDLFRHILNYYRTGKLHFPKN--E GI 1763619

105 YTLLYEEA-KYFQLQPMLLEMERWKQDRET 4094907
111 CLTSYDEELAFFGIMPDVIGDCCYEDYRDR GI 116443
114 CVSSFEDELTFFGIKGFNINNCCWDDYHDK GI 1763619
```

FIGURE 2

```
15   - - - R K R L L E Q E K S L A G W A L V L A G T G I G L M V   518158
271  L G H R R A L F E K R K R L S D Y A L I F G M F G I V V M V   GI 2564072

42   L H A E M L W F G G C S W A L Y L F L V K C T I S I S T F L   518158
301  I E T E L S W G L Y S K D S M F S L A L K C L I S L S T I I   GI 2564072

72   L L C L I V A F H A K E V Q L F M T D N G L R D W R V A L T   518158
331  L L G L I I A Y H T R E V Q L F V I D N G A D D W R I A M T   GI 2564072

102  G R Q A A Q I V L E L V V C G L H P A P V R - - - - - - - -   518158
361  Y E R I L Y I S L E M L V C A I H P I P G E Y K F F W T A R   GI 2564072

124  - - - - - - G P P C V Q D L G A P L T S P Q P W P G F L G Q   518158
391  L A F S Y T P S R A E A D V D I I L S I P M F L R L Y L I A   GI 2564072

148  G E A L L S L A M L - - - - - - - - - - - - - - - - - - - -   518158
421  R V M L L H S K L F T D A S S R S I G A L N K I N F N T R F   GI 2564072

158  - - - - - - - - - - - - - L L G L T L G L W L T T A W V L S   518158
451  V M K T L M T I C P G T V L L V F S I S L W I I A A W T V R   GI 2564072

175  V A E R - - Q A V N A T G H L S D T L W L I P I T F L T I G   518158
481  V C E R Y H D Q Q D V T S N F L G A M W L I S I T F L S I G   GI 2564072
```

FIGURE 3A

```
203  YGDVVPGTMWGKIVCLCTGVMGVCCTALLV  518158
511  YGDMVPHTYCGKGVCLLTGIMGAGCTALVV  GI 2564072

233  AVVARKLEFNKAEKHVHNFMMDIQYTKEMK  518158
541  AVVARKLELTKAEKHVHNFMMDTQLTKRIK  GI 2564072

263  ESAARVLQEAWMFYKHTR--RKESHA-ARR  518158
571  NAAANVLRETWLIYKHTKLLKKIDHAKVRK  GI 2564072

290  HQRKLLAAINAFRQVRLKHRKLREQVNSMV  518158
601  HQRKFLQAIHQLRGVKMEQRKLSDQANTLV  GI 2564072

320  DISKMHMILYDLQQNLSSSHRALEKQIDTL  518158
631  DLSKMQNVMYDLITELNDRSEDLEKQIGSL  GI 2564072

350  AGKLDALT-------------------ELL  518158
661  ESKLEHLTASFNSLPLLIADTLRQQQQQLL  GI 2564072

361  STALGPRQLP--------EPSQQSK       518158
691  TAFVEARGISVAVGTSHAPPSDSPIGISST  GI 2564072

377                                 518158
721  SFPTPYTSSSSC                    GI 2564072
```

FIGURE 3B

```
1    M V - - F G E F F - - - - - - - - - - - - - - - - - - - - -   602926
1    M N G D I G A W I S C A R T A G I G W V P I S S K E - P S A   GI 1763619
1    M A S - V A A W L P F A R A A A I G W V P I A T H P L P P P   GI 116443

8    - - - - - - - H R P G Q D E E L V N L N V G G F K Q S V D Q   602926
30   Y L N K Q V C N E N E K N N A K L T I N V S G R R Y Q T Y S   GI 1763619
30   P M P K - - - D R R K T D D E K L L I N V S G R R F E T W R   GI 116443

31   S T L L R F P H T R L G K L L T C H S E E A I L E L C D D Y   602926
60   H T L R K F K E T - - - - L L G S Q E R D Y F - - - - - - Y   GI 1763619
57   N T L E K Y P D T - - - - L L G S N E R E F F - - - - - - Y   GI 116443

61   S V A D K E Y Y F D R N P S S F R Y V L N F Y Y T G K L H V   602926
80   D E S L E E Y Y F D R D P D L F R H I L N Y Y R T G K L H F   GI 1763619
77   D E D C K E Y F F D R D P D I F R H I L N Y Y R T G K L H Y   GI 116443

91   M E E L C V F S F C Q E I E Y W G I N E L F I D S C C S N R   602926
110  P K N E C V S S F E D E L T F F G I K G F N I N N C C W D D   GI 1763619
107  P K H E C L T S Y D E E L A F F G I M P D V I G D C C Y E D   GI 116443

121  Y Q E R K E E N H E K D W D Q K S H D V S T D S S F E E S S   602926
140  Y H D K K R E C T E R L N E S D V M L T S S E I N E K S D T   GI 1763619
137  Y R D R K R E N A E R L M D - - - - - - - - - - - - - - D K   GI 116443
```

FIGURE 4A

```
151 L F E K E L E K F D T L R F G Q L R K K I W I R M E N P A -  602926
170 M G I D V Q M N N H Q A K - - N F R Q K V H G L F E N P Q S  GI 1763619
153 L S E N G D Q N L Q Q L T - - N M R Q K M W R A F E N P H T  GI 116443

180 - - - - - - - Y C L S A K L I A I S S L S V V L A S I V A M  602926
198 T F L A R I L Y Y I T G F F I A V S V G S T I I E T I D - -  GI 1763619
181 S T S A L V F Y Y V T G F F I A V S V M A N V V E T V P - -  GI 116443

203 C V H S M S E F Q N E D - G E V D D P V L E G V E I A C I A  602926
226 C S A N R - - - - - - P C G E V Y N K I F F N I E A V C V V  GI 1763619
209 C G H R P G R A G T L P C G E R Y K I V F F C L D T A C V M  GI 116443

232 W F T G E L A V R L A A A P C Q K K F W K N P L N I I D F V  602926
250 V F T I E Y L A R L Y S A P C R F R H A R I S L S I I D V I  GI 1763619
239 I F T A E Y L L R L F A A P D R C K F V R S V M S I I D V V  GI 116443

262 S I I P F Y A T L A V D T K E E E S E D I E N M G K V V Q I  602926
280 A I L P F Y I G L A M T - K T S I S G A F V S - - - - - - -  GI 1763619
269 A I M P Y Y I G L G I T D N D D V S G A F V T - - - - - - -  GI 116443

292 L R L M R I F R I L K L A R H S V G L R S L G A T L R H S Y  602926
302 L R V F R V F R I F K F S R H S K G L R I L G S T L T S C A  GI 1763619
292 L R V F R V F R I F K F S R H S Q G L R I L G Y T L K S C A  GI 116443
```

FIGURE 4B

```
322 HEVGLLLLFLSVGISIFSVLIYSVEKDDHT 602926
332 SELGFLLFSLSMAIIIIFATVVFYVEKDVND GI 1763619
322 SELGFLVFSLAMAIIIFATVMFYAEKNVNG GI 116443

352 SSLTSIPICWWWATISMTTVGYGDTHPVTL 602926
362 SDFTSIPASFWYTIVTMTTLGYGDMVPKTI GI 1763619
352 TNFTSIPAAFWYTIVTMTTLGYGDMVPETI GI 116443

382 AGKLIASTCIICGILVVALPITIIFNKFSK 602926
392 PGKLVGSICSLSGVLVIALPVPVIVSNFSR GI 1763619
382 AGKIVGGVCSLSGVLVIALPVPVIVSNFSR GI 116443

412 YYQKQKDIDVDQCSEDAPEKCHELPYFNIR 602926
422 IYLQNQRADKRRANQKLR------NKCEEK GI 1763619
412 IYHQNQRADKRKAQRKARLARIRIAKASSG GI 116443

442 DIYAQRMHAFITSLSSVGIVVSDPDSTDAS 602926
446 EE---KKK--ESSSETVTRFIISNQMY--- GI 1763619
442 AAFVSKKKAAEARWAAQESGIELDDNYRDE GI 116443

472 SIEDNEDICNTTSLENCTAK 602926
468 TIFSMKFAL----------TR GI 1763619
472 DIFELQHHHLLRCLEK-TTM GI 116443
```

FIGURE 4C

```
1    MGSGHCLRSTRGSKMVSWSVIAKIQEILQR  922119
1    MAG---------------SVLENIQSVLQK  GI 2350843

31   KMVREFLAEFMSTYVMMVFGLGSVAHMVLN  922119
16   TWVREFLAEFLNTYVLMVFGLGSVAHMVLG  GI 2350843

61   KKYGSYLGVNLGFGFGVTMGVHVAGRISGA  922119
46   ERLGSYLGVNLGFGFGVTMGIHVAGGISGA  GI 2350843

91   HMNAAVTFANCALGRVPWRKFPVYVLGQFL  922119
76   HMNPAVTFTNCALGRMAGRKFPIYVLGQFL  GI 2350843

121  GSFLAAATIYSLFYTAILHFSGGQLMVTGP  922119
106  GSFLAAATTYLIFYGAINHYAGETLLVTGP  GI 2350843

151  VATAGIFATYLPDHMTLWRGFLNEAWLTGM  922119
136  KSTANIFATYLPEHMTLWRGFVDEVFVTGM  GI 2350843

181  LQLCLFAITDQENNPALPGTEALVIGILVV  922119
166  LQLCIFAITDKLNSPALQGTEPLMIGILVC  GI 2350843

211  IIGVSLGMNTGYAINPSRDLPPRIFTFIAG  922119
196  VLGVSLGMNTGYAINPSRDLPPRFFTFIAG  GI 2350843
```

FIGURE 5A

```
241 W G K Q V F S N G E N W W W V P V V A P L L G A Y L G G I I 922119
226 W G K K V F S A G N N W W W V P V V A P L L G A Y L G G I V GI 2350843

271 Y L V F I G S T I P R E P L K L E D S V A Y E D H G I T V L 922119
256 Y L G L I H A G I P P Q - - - - - - - - - - - - - - - - - - GI 2350843

301 P K M G S H E P T I S P L T P V S V S P A N R S S V H P A P 922119
268 - - - G S                                                   GI 2350843

331 P L H E S M A L E H F                                       922119
269                                                             GI 2350843
```

FIGURE 5B

```
1    MNWAFLQGLLSGVNKYSTVLSRIWLSVVFI  2731369
1    MNWGFLQGILSGVNKYSTALGRIWLSVVFI  GI 192647

31   FRVLVYVVAAEEVWDDEQKDFDCNTKQPGC  2731369
31   FRVLVYVVAAEEVWDDDQKDFICNTKQPGC  GI 192647

61   TNVCYDNYFPISNIRLWALQLILVTCPSLL  2731369
61   PNVCYDEFFPVSHVRLWALQLILVTCPSLL  GI 192647

91   VVMHVAYREERERKHHLKHGPNAPSLYDNL  2731369
91   VVMHVAYREERERKHRLKHGPNAPALYSNL  GI 192647

121  SKKRGGLWWTYLLSLIFKAAVDAGFLYIFH  2731369
121  SKKRGGLWWTYLLSLIFKAAVDSGFLYIFH  GI 192647

151  RLYKDYDMPRVVACSVEPCPHTVDCYISRP  2731369
151  CIYKDYDMPRVVACSVTPCPHTVDCYIARP  GI 192647

181  TEKKVFTYFMVTTAAICILLNLSEVFYLVG  2731369
181  TEKKVFTYFMVVTAAICILLNLSEVVYLVG  GI 192647
```

FIGURE 6A

211	K R C M E I F G P R H R R P R C R E C L P D T C P P Y V L S	2731369
211	K R C M E V F R P R R R K A S R R H Q L P D T C P P Y V I S	GI 192647
241	Q G G H P E D G N S V L M K A G S A P V D A G G Y P	2731369
241	K G G H P Q D E S V I L T K A G M A T V D A G V Y P	GI 192647

FIGURE 6B

```
1    MFIWTSGRTSSSYRHDEKRNIYQKIRDHDL  2069907
1    MV----------------------------  g1055345

31   LDKRKTVTALKAGEDRAILLGLAMMVCSIM  2069907
3    ---KKLVMAQKRGETRALCLGVTMVVCAVI  g1055345

61   MYFLLGITLLRSYMQSVWTEESQCTLLNAS  2069907
30   TYYILVTTVLPLYQKSVWTQESKCHLIETN  g1055345

91   ITETFNCSFSCGPDCWKLSQYPCPQVYVNL  2069907
60   IRDQEE------LKGKKVPQYPC--LWVNV  g1055345

121  TSSGEKLLLYHTEETIKINQKCSYIPKCGK  2069907
82   SAAGRWAVLYHTEDTRDQNQQCSYIPGSVD  g1055345

151  NFEESMSLVNVVMENFRKYQHFSCYSDPEG  2069907
112  NYQTARADVEKVRAKFQEQQVFYCFSAPRG  g1055345

181  NQKSVILTKLYSSNVLFHSLFWPTCMMAGG  2069907
142  NETSVLFQRLYGPQALLFSLFWPTFLLTGG  g1055345

211  VAIVAMVKLTQYLSLLCERIQRINR       2069907
172  LLIIAMVKSNQYLSILAAQ------K      g1055345
```

FIGURE 7

```
1    M A E N H C E L L S P A R G G I G A G L G G G L C R R C S A  2243917
1    M S T V F I N - - - - - - - - - - - - - - - - - S R K S P N  g3292929

31   G L G A L A Q R P G S V S K W V R L N V G G T Y F L T T R Q  2243917
14   V L - - - - - - K K Q G T D Q W V K L N V G G T Y F L T T K T  g3292929

61   T L C R D P K S F L Y R L C Q A D P D L D S D K D E T G A Y  2243917
39   T L S R D P N S F L S R L I Q E D C D L I S D R D E T G A Y  g3292929

91   L I D R D P T Y F G P V L N Y L R H G K L V I N K D L A E E  2243917
69   L I D R D P K Y F A P V L N Y L R H G K L V L D - G V S E E  g3292929

121  G V L E E A E F Y N I T S L I K L V K D K I R E R D S K T S  2243917
98   G V L E E A E F Y N V T Q L I A L L K E C I L H R D Q R - P  g3292929

151  Q V P V K H V Y R V L Q C Q E E E L T Q M V S T M S D G W K  2243917
127  Q T D K K R V Y R V L Q C R E Q E L T Q M I S T L S D G W R  g3292929

181  F E Q L V S I G S S Y N Y G N E D Q A E F L C V V S K E L H  2243917
157  F E Q L I S M Q Y T - N Y G P F E N N E F L C V V S K E - C  g3292929

211  N T P Y G T A S E P S E K A K I L Q E R G S R - - - - - - -  2243917
185  G T T A G R E L E L N D R A K V L Q Q K G S R I N T I S H S  g3292929
```

FIGURE 8A

```
234 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 2243917
215 A T P T Q H Q L D A A K E A R A T A T A T S N T T N H T R S g3292929

234 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 2243917
245 D Q T Q P Q A Q I T H Q D Q P E S P K Q S P Q G D Y A S F A g3292929

234 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - 2243917
275 F E T K L T G T T A I R F S P L W P F C A L Y E V C A G V H g3292929

234 - - - M                                                     2243917
305 V F N L                                                     g3292929
```

FIGURE 8B

```
1    M V F - - T Q A P A E I M G H L R I R S L L A R Q C L A E F   2597476
1    M Q P E G A E K G K S F K Q R L V L K S S L A K E T L S E F   g2887407

29   L G V F V L M L L T Q G A V A Q A V T S G E T K G N F F T M   2597476
31   L G T F I L I V L G C G C V A Q A I L S R G R F G G V I T I   g2887407

59   F L A G S L A V T I A I Y V G G N V S G A H L N P A F S L A   2597476
61   N V G F S M A V A M A I Y V A G G V S G G H I N P A V S L A   g2887407

89   M C I V G R L P W V K L P I Y I L V Q L L S A F C A S G A T   2597476
91   M C L F G R M K W F K L P F Y V G A Q F L G A F V G A A T V   g2887407

119  Y V L Y H D A L Q N Y T G G N L T V T G P K E T A S I F A T   2597476
121  F G I Y Y D G L M S F A G G K L L I V G E N A T A H I F A T   g2887407

149  Y P A P Y L S L N N G F L D Q V L G T G M L I V G L L A I L   2597476
151  Y P A P Y L S L A N A F A D Q V V A T M I L L I I V F A I F   g2887407

179  D R R N K G V P A G L E P V V V G M L I L A L G L S M G A N   2597476
181  D S R N L G A P R G L E P I A I G L L I I V I A S S L G L N   g2887407

209  C G I P L N P A R D L G P R L F T Y V A G W G P E V F S A G   2597476
211  S G C A M N P A R D L S P R L F T A L A G W G F E V F R A G   g2887407
```

FIGURE 9A

```
239 NGWWWVPVVAPLVGATVGTATYQLLVALHH 2597476
241 NNFWWIPVVGPLVGAVIGGLIYVLVIEIHH g2887407

269 PEGPEPAQDLVSAQHKASELETPASAQMLE 2597476
271 PE-PDSVFKAEQSEDKPEKYE--------LS g2887407

299 CKL                            2597476
293 VIM                            g2887407
```

FIGURE 9B

SEQUENCE LISTING

<110> INCYTE PHARMACEUTICALS, INC.
AU-YOUNG, Janice
BANDMAN, Olga
TANG, Y. Tom
REDDY, Roopa
HILLMAN, Jennifer L.
YUE, Henry
LAL, Preeti
CORLEY, Neil C.
GUEGLER, Karl J.
GORGONE, Gina
BAUGHN, Mariah R.
AZIMZAI, Yalda

<120> HUMAN MEMBRANE CHANNEL PROTEINS

<130> PF-0589 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/145,815; unassigned; 09/191,283; unassigned; 09/208,821; unassigned
09/237,506; unassigned; 09/247;891; unassigned
<151> 1998-09-02; 1998-09-02; 1998-11-12; 1998-11-12; 1998-12-09; 1998-12-09
1999-01-26; 1999-01-26; 1999-02-10; 1999-02-10

<160> 45

<170> PERL Program

<210> 1
<211> 724
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1568324CD1

<400> 1
```
Met Ser Phe Glu Ser Ile Ser Ser Leu Pro Glu Val Glu Pro Asp
  1               5                  10                  15
Pro Glu Ala Gly Ser Glu Gln Glu Val Phe Ser Ala Val Glu Gly
                 20                  25                  30
Pro Ser Ala Glu Glu Thr Pro Ser Asp Thr Glu Ser Pro Glu Val
                 35                  40                  45
Leu Glu Thr Gln Leu Asp Ala His Gln Gly Leu Leu Gly Met Asp
                 50                  55                  60
Pro Pro Gly Asp Met Val Asp Phe Val Ala Ala Glu Ser Thr Glu
                 65                  70                  75
Asp Leu Lys Ala Leu Ser Ser Glu Glu Glu Glu Glu Met Gly Gly
                 80                  85                  90
Ala Ala Gln Glu Pro Glu Ser Leu Leu Pro Pro Ser Val Leu Asp
```

```
                        95                           100                          105
Gln Ala Ser Val Ile Ala Glu Arg Phe Val Ser Ser Phe Ser Arg
                110                 115                 120
Arg Ser Ser Val Ala Gln Glu Asp Ser Lys Ser Ser Gly Phe Gly
                125                 130                 135
Ser Pro Arg Leu Val Ser Arg Ser Ser Ser Val Leu Ser Leu Glu
                140                 145                 150
Gly Ser Glu Lys Gly Leu Ala Arg His Gly Ser Ala Thr Asp Ser
                155                 160                 165
Leu Ser Cys Gln Leu Ser Pro Glu Val Asp Ile Ser Val Gly Val
                170                 175                 180
Ala Thr Glu Asp Ser Pro Ser Val Asn Gly Met Glu Pro Pro Ser
                185                 190                 195
Pro Gly Cys Pro Val Glu Pro Asp Arg Ser Ser Cys Lys Lys Lys
                200                 205                 210
Glu Ser Ala Leu Ser Thr Arg Asp Arg Leu Leu Leu Asp Lys Ile
                215                 220                 225
Lys Ser Tyr Tyr Glu Asn Ala Glu His His Asp Ala Gly Phe Ser
                230                 235                 240
Val Arg Arg Arg Glu Ser Leu Ser Tyr Ile Pro Lys Gly Leu Val
                245                 250                 255
Arg Asn Ser Ile Ser Arg Phe Asn Ser Leu Pro Arg Pro Asp Pro
                260                 265                 270
Glu Pro Val Pro Pro Val Gly Ser Lys Arg Gln Val Gly Ser Arg
                275                 280                 285
Pro Thr Ser Trp Ala Leu Phe Glu Leu Pro Gly Pro Ser Gln Ala
                290                 295                 300
Val Lys Gly Asp Pro Pro Pro Ile Ser Asp Ala Glu Phe Arg Pro
                305                 310                 315
Ser Ser Glu Ile Val Lys Ile Trp Glu Gly Met Glu Ser Ser Gly
                320                 325                 330
Gly Ser Pro Gly Lys Gly Pro Gly Gln Gly Gln Ala Asn Gly Phe
                335                 340                 345
Asp Leu His Glu Pro Leu Phe Ile Leu Glu Glu His Glu Leu Gly
                350                 355                 360
Ala Ile Thr Glu Glu Ser Ala Thr Ala Ser Pro Glu Ser Ser Ser
                365                 370                 375
Pro Thr Glu Gly Arg Ser Pro Ala His Leu Ala Arg Glu Leu Lys
                380                 385                 390
Glu Leu Val Lys Glu Leu Ser Ser Ser Thr Gln Gly Glu Leu Val
                395                 400                 405
Ala Pro Leu His Pro Arg Ile Val Gln Leu Ser His Val Met Asp
                410                 415                 420
Ser His Val Ser Glu Arg Val Lys Asn Lys Val Tyr Gln Leu Ala
                425                 430                 435
Arg Gln Tyr Ser Leu Arg Ile Lys Ser Asn Lys Pro Val Met Ala
                440                 445                 450
Arg Pro Pro Leu Gln Trp Glu Lys Val Ala Pro Glu Arg Asp Gly
                455                 460                 465
Lys Ser Pro Thr Val Pro Cys Leu Gln Glu Glu Ala Gly Glu Pro
                470                 475                 480
Leu Gly Gly Lys Gly Lys Arg Lys Pro Val Leu Ser Leu Phe Asp
                485                 490                 495
Tyr Glu Gln Leu Met Ala Gln Glu His Ser Pro Pro Lys Pro Ser
                500                 505                 510
Ser Ala Gly Glu Met Ser Pro Gln Arg Phe Phe Phe Asn Pro Pro
```

515 520 525
Ala Val Ser Gln Arg Thr Thr Ser Pro Gly Gly Arg Pro Ser Ala
530 535 540
Arg Ser Pro Leu Ser Pro Thr Glu Thr Phe Ser Trp Pro Asp Val
545 550 555
Arg Glu Leu Cys Ser Lys Tyr Ala Ser Arg Asp Glu Ala Arg Arg
560 565 570
Ala Gly Gly Gly Arg Pro Arg Gly Pro Pro Val Asn Arg Ser His
575 580 585
Ser Val Pro Glu Asn Met Val Glu Pro Pro Leu Ser Gly Arg Val
590 595 600
Gly Arg Cys Arg Ser Leu Ser Thr Lys Arg Gly Arg Gly Gly Gly
605 610 615
Glu Ala Ala Gln Ser Pro Gly Pro Leu Pro Gln Ser Lys Pro Asp
620 625 630
Gly Gly Glu Thr Leu Tyr Val Thr Ala Asp Leu Thr Leu Glu Asp
635 640 645
Asn Arg Arg Val Ile Val Met Glu Lys Gly Pro Leu Pro Ser Pro
650 655 660
Thr Ala Gly Leu Glu Glu Ser Ser Gly Gln Gly Pro Ser Ser Pro
665 670 675
Val Ala Leu Leu Gly Gln Val Gln Asp Phe Gln Gln Ser Ala Glu
680 685 690
Cys Gln Pro Lys Glu Glu Gly Ser Arg Asp Pro Ala Asp Pro Ser
695 700 705
Gln Gln Gly Arg Val Arg Asn Leu Arg Glu Lys Phe Gln Ala Leu
710 715 720
Asn Ser Val Gly

<210> 2
<211> 257
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4094907CD1

<400> 2
Met Ser Arg Pro Leu Ile Thr Arg Ser Pro Ala Ser Pro Leu Asn
1 5 10 15
Asn Gln Gly Ile Pro Thr Pro Ala Gln Leu Thr Lys Ser Asn Ala
20 25 30
Pro Val His Ile Asp Val Gly Gly His Met Tyr Thr Ser Ser Leu
35 40 45
Ala Thr Leu Thr Lys Tyr Pro Glu Ser Arg Ile Gly Arg Leu Phe
50 55 60
Asp Gly Thr Glu Pro Ile Val Leu Asp Ser Leu Lys Gln His Tyr
65 70 75
Phe Ile Asp Arg Asp Gly Gln Met Phe Arg Tyr Ile Leu Asn Phe
80 85 90
Leu Arg Thr Ser Lys Leu Leu Ile Pro Asp Asp Phe Lys Asp Tyr
95 100 105
Thr Leu Leu Tyr Glu Glu Ala Lys Tyr Phe Gln Leu Gln Pro Met
110 115 120

Leu Leu Glu Met Glu Arg Trp Lys Gln Asp Arg Glu Thr Gly Arg
125 130 135
Phe Ser Arg Pro Cys Glu Cys Leu Val Val Arg Val Ala Pro Asp
140 145 150
Leu Gly Glu Arg Ile Thr Leu Ser Gly Asp Lys Ser Leu Ile Glu
155 160 165
Glu Val Phe Pro Glu Ile Gly Asp Val Met Cys Asn Ser Val Asn
170 175 180
Ala Gly Trp Asn His Asp Ser Thr His Val Ile Arg Phe Pro Leu
185 190 195
Asn Gly Tyr Cys His Leu Asn Ser Val Gln Val Leu Glu Arg Leu
200 205 210
Gln Gln Arg Gly Phe Glu Ile Val Gly Ser Cys Gly Gly Gly Val
215 220 225
Asp Ser Ser Gln Phe Ser Glu Tyr Val Leu Arg Arg Glu Leu Arg
230 235 240
Arg Thr Pro Arg Val Pro Ser Val Ile Arg Ile Lys Gln Glu Pro
245 250 255
Leu Asp

<210> 3
<211> 377
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 518158CD1

<400> 3
Met Gly Gly Asp Leu Val Leu Gly Leu Gly Ala Leu Arg Arg Arg
1 5 10 15
Lys Arg Leu Leu Glu Gln Glu Lys Ser Leu Ala Gly Trp Ala Leu
20 25 30
Val Leu Ala Gly Thr Gly Ile Gly Leu Met Val Leu His Ala Glu
35 40 45
Met Leu Trp Phe Gly Gly Cys Ser Trp Ala Leu Tyr Leu Phe Leu
50 55 60
Val Lys Cys Thr Ile Ser Ile Ser Thr Phe Leu Leu Leu Cys Leu
65 70 75
Ile Val Ala Phe His Ala Lys Glu Val Gln Leu Phe Met Thr Asp
80 85 90
Asn Gly Leu Arg Asp Trp Arg Val Ala Leu Thr Gly Arg Gln Ala
95 100 105
Ala Gln Ile Val Leu Glu Leu Val Val Cys Gly Leu His Pro Ala
110 115 120
Pro Val Arg Gly Pro Pro Cys Val Gln Asp Leu Gly Ala Pro Leu
125 130 135
Thr Ser Pro Gln Pro Trp Pro Gly Phe Leu Gly Gln Gly Glu Ala
140 145 150
Leu Leu Ser Leu Ala Met Leu Leu Leu Gly Leu Thr Leu Gly Leu
155 160 165
Trp Leu Thr Thr Ala Trp Val Leu Ser Val Ala Glu Arg Gln Ala
170 175 180
Val Asn Ala Thr Gly His Leu Ser Asp Thr Leu Trp Leu Ile Pro

```
                185                 190                 195
Ile Thr Phe Leu Thr Ile Gly Tyr Gly Asp Val Val Pro Gly Thr
                200                 205                 210
Met Trp Gly Lys Ile Val Cys Leu Cys Thr Gly Val Met Gly Val
                215                 220                 225
Cys Cys Thr Ala Leu Leu Val Ala Val Val Ala Arg Lys Leu Glu
                230                 235                 240
Phe Asn Lys Ala Glu Lys His Val His Asn Phe Met Met Asp Ile
                245                 250                 255
Gln Tyr Thr Lys Glu Met Lys Glu Ser Ala Ala Arg Val Leu Gln
                260                 265                 270
Glu Ala Trp Met Phe Tyr Lys His Thr Arg Arg Lys Glu Ser His
                275                 280                 285
Ala Ala Arg Arg His Gln Arg Lys Leu Leu Ala Ala Ile Asn Ala
                290                 295                 300
Phe Arg Gln Val Arg Leu Lys His Arg Lys Leu Arg Glu Gln Val
                305                 310                 315
Asn Ser Met Val Asp Ile Ser Lys Met His Met Ile Leu Tyr Asp
                320                 325                 330
Leu Gln Gln Asn Leu Ser Ser Ser His Arg Ala Leu Glu Lys Gln
                335                 340                 345
Ile Asp Thr Leu Ala Gly Lys Leu Asp Ala Leu Thr Glu Leu Leu
                350                 355                 360
Ser Thr Ala Leu Gly Pro Arg Gln Leu Pro Glu Pro Ser Gln Gln
                365                 370                 375
Ser Lys


<210> 4
<211> 491
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 602926CD1

<400> 4
Met Val Phe Gly Glu Phe Phe His Arg Pro Gly Gln Asp Glu Glu
  1               5                  10                  15
Leu Val Asn Leu Asn Val Gly Gly Phe Lys Gln Ser Val Asp Gln
                 20                  25                  30
Ser Thr Leu Leu Arg Phe Pro His Thr Arg Leu Gly Lys Leu Leu
                 35                  40                  45
Thr Cys His Ser Glu Glu Ala Ile Leu Glu Leu Cys Asp Asp Tyr
                 50                  55                  60
Ser Val Ala Asp Lys Glu Tyr Tyr Phe Asp Arg Asn Pro Ser Ser
                 65                  70                  75
Phe Arg Tyr Val Leu Asn Phe Tyr Tyr Thr Gly Lys Leu His Val
                 80                  85                  90
Met Glu Glu Leu Cys Val Phe Ser Phe Cys Gln Glu Ile Glu Tyr
                 95                 100                 105
Trp Gly Ile Asn Glu Leu Phe Ile Asp Ser Cys Cys Ser Asn Arg
                110                 115                 120
Tyr Gln Glu Arg Lys Glu Glu Asn His Glu Lys Asp Trp Asp Gln
                125                 130                 135
```

Lys Ser His Asp Val Ser Thr Asp Ser Ser Phe Glu Glu Ser Ser
140 145 150
Leu Phe Glu Lys Glu Leu Glu Lys Phe Asp Thr Leu Arg Phe Gly
155 160 165
Gln Leu Arg Lys Lys Ile Trp Ile Arg Met Glu Asn Pro Ala Tyr
170 175 180
Cys Leu Ser Ala Lys Leu Ile Ala Ile Ser Ser Leu Ser Val Val
185 190 195
Leu Ala Ser Ile Val Ala Met Cys Val His Ser Met Ser Glu Phe
200 205 210
Gln Asn Glu Asp Gly Glu Val Asp Asp Pro Val Leu Glu Gly Val
215 220 225
Glu Ile Ala Cys Ile Ala Trp Phe Thr Gly Glu Leu Ala Val Arg
230 235 240
Leu Ala Ala Ala Pro Cys Gln Lys Lys Phe Trp Lys Asn Pro Leu
245 250 255
Asn Ile Ile Asp Phe Val Ser Ile Ile Pro Phe Tyr Ala Thr Leu
260 265 270
Ala Val Asp Thr Lys Glu Glu Glu Ser Glu Asp Ile Glu Asn Met
275 280 285
Gly Lys Val Val Gln Ile Leu Arg Leu Met Arg Ile Phe Arg Ile
290 295 300
Leu Lys Leu Ala Arg His Ser Val Gly Leu Arg Ser Leu Gly Ala
305 310 315
Thr Leu Arg His Ser Tyr His Glu Val Gly Leu Leu Leu Leu Phe
320 325 330
Leu Ser Val Gly Ile Ser Ile Phe Ser Val Leu Ile Tyr Ser Val
335 340 345
Glu Lys Asp Asp His Thr Ser Ser Leu Thr Ser Ile Pro Ile Cys
350 355 360
Trp Trp Trp Ala Thr Ile Ser Met Thr Thr Val Gly Tyr Gly Asp
365 370 375
Thr His Pro Val Thr Leu Ala Gly Lys Leu Ile Ala Ser Thr Cys
380 385 390
Ile Ile Cys Gly Ile Leu Val Val Ala Leu Pro Ile Thr Ile Ile
395 400 405
Phe Asn Lys Phe Ser Lys Tyr Tyr Gln Lys Gln Lys Asp Ile Asp
410 415 420
Val Asp Gln Cys Ser Glu Asp Ala Pro Glu Lys Cys His Glu Leu
425 430 435
Pro Tyr Phe Asn Ile Arg Asp Ile Tyr Ala Gln Arg Met His Ala
440 445 450
Phe Ile Thr Ser Leu Ser Ser Val Gly Ile Val Val Ser Asp Pro
455 460 465
Asp Ser Thr Asp Ala Ser Ser Ile Glu Asp Asn Glu Asp Ile Cys
470 475 480
Asn Thr Thr Ser Leu Glu Asn Cys Thr Ala Lys
485 490

<210> 5
<211> 341
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 922119CD1

<400> 5
Met Gly Ser Gly His Cys Leu Arg Ser Thr Arg Gly Ser Lys Met
1 5 10 15
Val Ser Trp Ser Val Ile Ala Lys Ile Gln Glu Ile Leu Gln Arg
20 25 30
Lys Met Val Arg Glu Phe Leu Ala Glu Phe Met Ser Thr Tyr Val
35 40 45
Met Met Val Phe Gly Leu Gly Ser Val Ala His Met Val Leu Asn
50 55 60
Lys Lys Tyr Gly Ser Tyr Leu Gly Val Asn Leu Gly Phe Gly Phe
65 70 75
Gly Val Thr Met Gly Val His Val Ala Gly Arg Ile Ser Gly Ala
80 85 90
His Met Asn Ala Ala Val Thr Phe Ala Asn Cys Ala Leu Gly Arg
95 100 105
Val Pro Trp Arg Lys Phe Pro Val Tyr Val Leu Gly Gln Phe Leu
110 115 120
Gly Ser Phe Leu Ala Ala Ala Thr Ile Tyr Ser Leu Phe Tyr Thr
125 130 135
Ala Ile Leu His Phe Ser Gly Gly Gln Leu Met Val Thr Gly Pro
140 145 150
Val Ala Thr Ala Gly Ile Phe Ala Thr Tyr Leu Pro Asp His Met
155 160 165
Thr Leu Trp Arg Gly Phe Leu Asn Glu Ala Trp Leu Thr Gly Met
170 175 180
Leu Gln Leu Cys Leu Phe Ala Ile Thr Asp Gln Glu Asn Asn Pro
185 190 195
Ala Leu Pro Gly Thr Glu Ala Leu Val Ile Gly Ile Leu Val Val
200 205 210
Ile Ile Gly Val Ser Leu Gly Met Asn Thr Gly Tyr Ala Ile Asn
215 220 225
Pro Ser Arg Asp Leu Pro Pro Arg Ile Phe Thr Phe Ile Ala Gly
230 235 240
Trp Gly Lys Gln Val Phe Ser Asn Gly Glu Asn Trp Trp Trp Val
245 250 255
Pro Val Val Ala Pro Leu Leu Gly Ala Tyr Leu Gly Gly Ile Ile
260 265 270
Tyr Leu Val Phe Ile Gly Ser Thr Ile Pro Arg Glu Pro Leu Lys
275 280 285
Leu Glu Asp Ser Val Ala Tyr Glu Asp His Gly Ile Thr Val Leu
290 295 300
Pro Lys Met Gly Ser His Glu Pro Thr Ile Ser Pro Leu Thr Pro
305 310 315
Val Ser Val Ser Pro Ala Asn Arg Ser Ser Val His Pro Ala Pro
320 325 330
Pro Leu His Glu Ser Met Ala Leu Glu His Phe
335 340

<210> 6
<211> 476
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2666782CD1

<400> 6
```
Met Gly Ile Lys Phe Leu Glu Val Ile Lys Pro Phe Cys Ala Val
 1               5                  10                  15
Leu Pro Glu Ile Gln Lys Pro Glu Arg Lys Ile Gln Phe Arg Glu
                20                  25                  30
Lys Val Leu Trp Thr Ala Ile Thr Leu Phe Ile Phe Leu Val Cys
                35                  40                  45
Cys Gln Ile Pro Leu Phe Gly Ile Met Ser Ser Asp Ser Ala Asp
                50                  55                  60
Pro Phe Tyr Trp Met Arg Val Ile Leu Ala Ser Asn Arg Gly Thr
                65                  70                  75
Leu Met Glu Leu Gly Ile Ser Pro Ile Val Thr Ser Gly Leu Ile
                80                  85                  90
Met Gln Leu Leu Ala Gly Ala Lys Ile Ile Glu Val Gly Asp Thr
                95                  100                 105
Pro Lys Asp Arg Ala Leu Phe Asn Gly Ala Gln Lys Leu Phe Gly
                110                 115                 120
Met Ile Ile Thr Ile Gly Gln Ala Ile Val Tyr Val Met Thr Gly
                125                 130                 135
Met Tyr Gly Asp Pro Ala Glu Met Gly Ala Gly Ile Cys Leu Leu
                140                 145                 150
Ile Ile Ile Gln Leu Phe Val Thr Ser Leu Ile Val Leu Leu Leu
                155                 160                 165
Asp Glu Leu Leu Gln Thr Gly Tyr Ser Leu Gly Ser Gly Ile Ser
                170                 175                 180
Leu Val Ile Ala Thr Asn Ile Cys Glu Thr Ile Val Trp Lys Ala
                185                 190                 195
Phe Ser Pro Thr Thr Ile Asn Thr Gly Arg Gly Thr Glu Phe Glu
                200                 205                 210
Gly Ala Val Ile Ala Leu Phe His Leu Leu Ala Thr Arg Thr Asp
                215                 220                 225
Lys Val Arg Ala Leu Arg Glu Ala Phe Tyr Arg Gln Asn Leu Pro
                230                 235                 240
Asn Leu Met Asn Leu Ile Ala Thr Val Phe Val Phe Ala Val Val
                245                 250                 255
Ile Tyr Phe Gln Gly Phe Arg Val Asp Leu Pro Ile Lys Ser Ala
                260                 265                 270
Arg Tyr Arg Gly Gln Tyr Ser Ser Tyr Pro Ile Lys Leu Phe Tyr
                275                 280                 285
Thr Ser Asn Ile Pro Ile Ile Leu Gln Ser Ala Leu Val Ser Asn
                290                 295                 300
Leu Tyr Val Ile Ser Gln Met Leu Ser Val Arg Phe Ser Gly Asn
                305                 310                 315
Phe Leu Val Asn Leu Leu Gly Gln Trp Ala Asp Val Ser Gly Gly
                320                 325                 330
Gly Pro Ala Arg Ser Tyr Pro Val Gly Gly Leu Cys Tyr Tyr Leu
                335                 340                 345
Ser Pro Pro Glu Ser Met Gly Ala Ile Phe Glu Asp Pro Val His
                350                 355                 360
Val Val Val Tyr Ile Ile Phe Met Leu Gly Ser Cys Ala Phe Phe
                365                 370                 375
Ser Lys Thr Trp Ile Glu Val Ser Gly Ser Ser Ala Lys Asp Val
```

```
                    380                 385                 390
Ala Lys Gln Leu Lys Glu Gln Gln Met Val Met Arg Gly His Arg
                    395                 400                 405
Asp Thr Ser Met Val His Glu Leu Asn Arg Tyr Ile Pro Thr Ala
                    410                 415                 420
Ala Ala Phe Gly Gly Leu Cys Ile Gly Ala Leu Ser Val Leu Ala
                    425                 430                 435
Asp Phe Leu Gly Ala Ile Gly Ser Gly Thr Gly Ile Leu Leu Ala
                    440                 445                 450
Val Thr Ile Ile Tyr Gln Tyr Phe Glu Ile Phe Val Lys Glu Gln
                    455                 460                 465
Ala Glu Val Gly Gly Met Gly Ala Leu Phe Phe
                    470                 475
```

<210> 7
<211> 266
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2731369CD1

<400> 7

```
Met Asn Trp Ala Phe Leu Gln Gly Leu Leu Ser Gly Val Asn Lys
  1               5                  10                  15
Tyr Ser Thr Val Leu Ser Arg Ile Trp Leu Ser Val Val Phe Ile
                 20                  25                  30
Phe Arg Val Leu Val Tyr Val Val Ala Ala Glu Glu Val Trp Asp
                 35                  40                  45
Asp Glu Gln Lys Asp Phe Asp Cys Asn Thr Lys Gln Pro Gly Cys
                 50                  55                  60
Thr Asn Val Cys Tyr Asp Asn Tyr Phe Pro Ile Ser Asn Ile Arg
                 65                  70                  75
Leu Trp Ala Leu Gln Leu Ile Leu Val Thr Cys Pro Ser Leu Leu
                 80                  85                  90
Val Val Met His Val Ala Tyr Arg Glu Glu Arg Glu Arg Lys His
                 95                 100                 105
His Leu Lys His Gly Pro Asn Ala Pro Ser Leu Tyr Asp Asn Leu
                110                 115                 120
Ser Lys Lys Arg Gly Gly Leu Trp Trp Thr Tyr Leu Leu Ser Leu
                125                 130                 135
Ile Phe Lys Ala Ala Val Asp Ala Gly Phe Leu Tyr Ile Phe His
                140                 145                 150
Arg Leu Tyr Lys Asp Tyr Asp Met Pro Arg Val Val Ala Cys Ser
                155                 160                 165
Val Glu Pro Cys Pro His Thr Val Asp Cys Tyr Ile Ser Arg Pro
                170                 175                 180
Thr Glu Lys Lys Val Phe Thr Tyr Phe Met Val Thr Thr Ala Ala
                185                 190                 195
Ile Cys Ile Leu Leu Asn Leu Ser Glu Val Phe Tyr Leu Val Gly
                200                 205                 210
Lys Arg Cys Met Glu Ile Phe Gly Pro Arg His Arg Arg Pro Arg
                215                 220                 225
Cys Arg Glu Cys Leu Pro Asp Thr Cys Pro Pro Tyr Val Leu Ser
```

230 235 240
Gln Gly Gly His Pro Glu Asp Gly Asn Ser Val Leu Met Lys Ala
245 250 255
Gly Ser Ala Pro Val Asp Ala Gly Gly Tyr Pro
260 265

<210> 8
<211> 182
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1375415CD1

<400> 8
Met Ala Glu Phe Pro Ser Lys Val Ser Thr Arg Thr Ser Ser Pro
1 5 10 15
Ala Gln Gly Ala Glu Ala Ser Val Ser Ala Leu Arg Pro Asp Leu
20 25 30
Gly Phe Val Arg Ser Arg Leu Gly Ala Leu Met Leu Leu Gln Leu
35 40 45
Val Leu Gly Leu Leu Val Trp Ala Leu Ile Ala Asp Thr Pro Tyr
50 55 60
His Leu Tyr Pro Ala Tyr Gly Trp Val Met Phe Val Ala Val Phe
65 70 75
Leu Trp Leu Val Thr Ile Val Leu Phe Asn Leu Tyr Leu Phe Gln
80 85 90
Leu His Met Lys Leu Tyr Met Val Pro Trp Pro Leu Val Leu Met
95 100 105
Ile Phe Asn Ile Ser Ala Thr Val Leu Tyr Ile Thr Ala Phe Ile
110 115 120
Ala Cys Ser Ala Ala Val Asp Leu Thr Ser Leu Arg Gly Thr Arg
125 130 135
Pro Tyr Asn Gln Arg Ala Ala Ala Ser Phe Phe Ala Cys Leu Val
140 145 150
Met Ile Ala Tyr Gly Val Ser Ala Phe Phe Ser Tyr Gln Ala Trp
155 160 165
Arg Gly Val Gly Ser Asn Ala Ala Thr Ser Gln Met Ala Gly Gly
170 175 180
Tyr Ala

<210> 9
<211> 942
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2733282CD1

<400> 9
Met Thr Gln Arg Ser Ile Ala Gly Pro Ile Cys Asn Leu Lys Phe
1 5 10 15

```
Val Thr Leu Leu Val Ala Leu Ser Ser Glu Leu Pro Phe Leu Gly
                20                  25                  30
Ala Gly Val Gln Leu Gln Asp Asn Gly Tyr Asn Gly Leu Leu Ile
                35                  40                  45
Ala Ile Asn Pro Gln Val Pro Glu Asn Gln Asn Leu Ile Ser Asn
                50                  55                  60
Ile Lys Glu Met Ile Thr Glu Ala Ser Phe Tyr Leu Phe Asn Ala
                65                  70                  75
Thr Lys Arg Arg Val Phe Phe Arg Asn Ile Lys Ile Leu Ile Pro
                80                  85                  90
Ala Thr Trp Lys Ala Asn Asn Asn Ser Lys Ile Lys Gln Glu Ser
                95                  100                 105
Tyr Glu Lys Ala Asn Val Ile Val Thr Asp Trp Tyr Gly Ala His
                110                 115                 120
Gly Asp Asp Pro Tyr Thr Leu Gln Tyr Arg Gly Cys Gly Lys Glu
                125                 130                 135
Gly Lys Tyr Ile His Phe Thr Pro Asn Phe Leu Leu Asn Asp Asn
                140                 145                 150
Leu Thr Ala Gly Tyr Gly Ser Arg Gly Arg Val Phe Val His Glu
                155                 160                 165
Trp Ala His Leu Arg Trp Gly Val Phe Asp Glu Tyr Asn Asn Asp
                170                 175                 180
Lys Pro Phe Tyr Ile Asn Gly Gln Asn Gln Ile Lys Val Thr Arg
                185                 190                 195
Cys Ser Ser Asp Ile Thr Gly Ile Phe Val Cys Glu Lys Gly Pro
                200                 205                 210
Cys Pro Gln Glu Asn Cys Ile Ile Ser Lys Leu Phe Lys Glu Gly
                215                 220                 225
Cys Thr Phe Ile Tyr Asn Ser Thr Gln Asn Ala Thr Ala Ser Ile
                230                 235                 240
Met Phe Met Gln Ser Tyr Leu Cys Gly Glu Ile Cys Asn Ala Ser
                245                 250                 255
Thr His Asn Gln Glu Ala Pro Asn Leu Gln Asn Gln Met Cys Ser
                260                 265                 270
Leu Arg Ser Ala Trp Asp Val Ile Thr Asp Ser Ala Asp Phe His
                275                 280                 285
His Ser Phe Pro Met Asn Gly Thr Glu Leu Pro Pro Pro Pro Thr
                290                 295                 300
Phe Ser Leu Val Glu Ala Gly Asp Lys Val Val Cys Leu Val Leu
                305                 310                 315
Asp Val Ser Ser Lys Met Ala Glu Ala Asp Arg Leu Leu Gln Leu
                320                 325                 330
Gln Gln Ala Ala Glu Phe Tyr Leu Met Gln Ile Val Glu Ile His
                335                 340                 345
Thr Phe Val Gly Ile Ala Ser Phe Asp Ser Lys Gly Glu Ile Arg
                350                 355                 360
Ala Gln Leu His Gln Ile Asn Ser Asn Asp Asp Arg Lys Leu Leu
                365                 370                 375
Val Ser Tyr Leu Pro Thr Thr Val Ser Ala Lys Thr Asp Ile Ser
                380                 385                 390
Ile Cys Ser Gly Leu Lys Lys Gly Phe Glu Val Val Glu Lys Leu
                395                 400                 405
Asn Gly Lys Ala Tyr Gly Ser Val Met Ile Leu Val Thr Ser Gly
                410                 415                 420
Asp Asp Lys Leu Leu Gly Asn Cys Leu Pro Thr Val Leu Ser Ser
                425                 430                 435
```

Gly Ser Thr Ile His Ser Ile Ala Leu Gly Ser Ser Ala Ala Pro
440 445 450
Asn Leu Glu Glu Leu Ser Arg Leu Thr Gly Gly Leu Lys Phe Phe
455 460 465
Val Pro Asp Ile Ser Asn Ser Asn Ser Met Ile Asp Ala Phe Ser
470 475 480
Arg Ile Ser Ser Gly Thr Gly Asp Ile Phe Gln Gln His Ile Gln
485 490 495
Leu Glu Ser Thr Gly Glu Asn Val Lys Pro His His Gln Leu Lys
500 505 510
Asn Thr Val Thr Val Asp Asn Thr Val Gly Asn Asp Thr Met Phe
515 520 525
Leu Val Thr Trp Gln Ala Ser Gly Pro Pro Glu Ile Ile Leu Phe
530 535 540
Asp Pro Asp Gly Arg Lys Tyr Tyr Thr Asn Asn Phe Ile Thr Asn
545 550 555
Leu Thr Phe Arg Thr Ala Ser Leu Trp Ile Pro Gly Thr Ala Lys
560 565 570
Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His His Ser Leu
575 580 585
Gln Ala Leu Lys Val Thr Val Thr Ser Arg Ala Ser Asn Ser Ala
590 595 600
Val Pro Pro Ala Thr Val Glu Ala Phe Val Glu Arg Asp Ser Leu
605 610 615
His Phe Pro His Pro Val Met Ile Tyr Ala Asn Val Lys Gln Gly
620 625 630
Phe Tyr Pro Ile Leu Asn Ala Thr Val Thr Ala Thr Val Glu Pro
635 640 645
Glu Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala
650 655 660
Gly Ala Asp Val Ile Lys Asn Asp Gly Ile Tyr Ser Arg Tyr Phe
665 670 675
Phe Ser Phe Ala Ala Asn Gly Arg Tyr Ser Leu Lys Val His Val
680 685 690
Asn His Ser Pro Ser Ile Ser Thr Pro Ala His Ser Ile Pro Gly
695 700 705
Ser His Ala Met Tyr Val Pro Gly Tyr Thr Ala Asn Gly Asn Ile
710 715 720
Gln Met Asn Ala Pro Arg Lys Ser Val Gly Arg Asn Glu Glu Glu
725 730 735
Arg Lys Trp Gly Phe Ser Arg Val Ser Ser Gly Gly Ser Phe Ser
740 745 750
Val Leu Gly Val Pro Ala Gly Pro His Pro Asp Val Phe Pro Pro
755 760 765
Cys Lys Ile Ile Asp Leu Glu Ala Val Lys Val Glu Glu Glu Leu
770 775 780
Thr Leu Ser Trp Thr Ala Pro Gly Glu Asp Phe Asp Gln Gly Gln
785 790 795
Ala Thr Ser Tyr Glu Ile Arg Met Ser Lys Ser Leu Gln Asn Ile
800 805 810
Gln Asp Asp Phe Asn Asn Ala Ile Leu Val Asn Thr Ser Lys Arg
815 820 825
Asn Pro Gln Gln Ala Gly Ile Arg Glu Ile Phe Thr Phe Ser Pro
830 835 840
Gln Ile Ser Thr Asn Gly Pro Glu His Gln Pro Asn Gly Glu Thr
845 850 855

His Glu Ser His Arg Ile Tyr Val Ala Ile Arg Ala Met Asp Arg
860 865 870
Asn Ser Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu
875 880 885
Phe Ile Pro Pro Asn Ser Asp Pro Val Pro Ala Arg Asp Tyr Leu
890 895 900
Ile Leu Lys Gly Val Leu Thr Ala Met Gly Leu Ile Gly Ile Ile
905 910 915
Cys Leu Ile Ile Val Val Thr His His Thr Leu Ser Arg Lys Lys
920 925 930
Arg Ala Asp Lys Lys Glu Asn Gly Thr Lys Leu Leu
935 940

<210> 10
<211> 519
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3148427CD1

<400> 10
Met Glu Glu Met Phe His Lys Lys Ser Glu Ala Val Arg Arg Leu
1 5 10 15
Val Glu Ala Ala Glu Glu Ala His Leu Lys His Glu Phe Asp Ala
20 25 30
Asp Leu Gln Tyr Glu Tyr Phe Asn Ala Val Leu Ile Asn Glu Arg
35 40 45
Asp Lys Asp Gly Asn Phe Leu Glu Leu Gly Lys Glu Phe Ile Leu
50 55 60
Ala Pro Asn Asp His Phe Asn Asn Leu Pro Val Asn Ile Ser Leu
65 70 75
Ser Asp Val Gln Val Pro Thr Asn Met Tyr Asn Lys Asp Pro Ala
80 85 90
Ile Val Asn Gly Val Tyr Trp Ser Glu Ser Leu Asn Lys Val Phe
95 100 105
Val Asp Asn Phe Asp Arg Asp Pro Ser Leu Ile Trp Gln Tyr Phe
110 115 120
Gly Ser Ala Lys Gly Phe Phe Arg Gln Tyr Pro Gly Ile Lys Trp
125 130 135
Glu Pro Asp Glu Asn Gly Val Ile Ala Phe Asp Cys Arg Asn Arg
140 145 150
Lys Trp Tyr Ile Gln Ala Ala Thr Ser Pro Lys Asp Val Val Ile
155 160 165
Leu Val Asp Val Ser Gly Ser Met Lys Gly Leu Arg Leu Thr Ile
170 175 180
Ala Lys Gln Thr Val Ser Ser Ile Leu Asp Thr Leu Gly Asp Asp
185 190 195
Asp Phe Phe Asn Ile Ile Ala Tyr Asn Glu Glu Leu His Tyr Val
200 205 210
Glu Pro Cys Leu Asn Gly Thr Leu Val Gln Ala Asp Arg Thr Asn
215 220 225
Lys Glu His Phe Arg Glu His Leu Asp Lys Leu Phe Ala Lys Gly
230 235 240

Ile Gly Met Leu Asp Ile Ala Leu Asn Glu Ala Phe Asn Ile Leu
245 250 255
Ser Asp Phe Asn His Thr Gly Gln Gly Ser Ile Cys Ser Gln Ala
260 265 270
Ile Met Leu Ile Thr Asp Gly Ala Val Asp Thr Tyr Asp Thr Ile
275 280 285
Phe Ala Lys Tyr Asn Trp Pro Asp Arg Lys Val Arg Ile Phe Thr
290 295 300
Tyr Leu Ile Gly Arg Glu Ala Ala Phe Ala Asp Asn Leu Lys Trp
305 310 315
Met Ala Cys Ala Asn Lys Gly Phe Phe Thr Gln Ile Ser Thr Leu
320 325 330
Ala Asp Val Gln Glu Asn Val Met Glu Tyr Leu His Val Leu Ser
335 340 345
Arg Pro Lys Val Ile Asp Gln Glu His Asp Val Val Trp Thr Glu
350 355 360
Ala Tyr Ile Asp Ser Thr Leu Pro Gln Ala Gln Lys Leu Thr Asp
365 370 375
Asp Gln Gly Pro Val Leu Met Thr Thr Val Ala Met Pro Val Phe
380 385 390
Ser Lys Gln Asn Glu Thr Arg Ser Lys Gly Ile Leu Leu Gly Val
395 400 405
Val Gly Thr Asp Val Pro Val Lys Glu Leu Leu Lys Thr Ile Pro
410 415 420
Lys Tyr Lys Leu Gly Ile His Gly Tyr Ala Phe Ala Ile Thr Asn
425 430 435
Asn Gly Tyr Ile Leu Thr His Pro Glu Leu Arg Leu Leu Tyr Glu
440 445 450
Glu Gly Lys Lys Arg Arg Lys Pro Asn Tyr Ser Ser Val Asp Leu
455 460 465
Ser Glu Val Glu Trp Glu Asp Arg Asp Asp Val Leu Arg Asn Ala
470 475 480
Met Val Asn Arg Lys Thr Gly Lys Phe Ser Met Glu Val Lys Lys
485 490 495
Thr Val Asp Lys Gly Val His Phe Ser Gln Thr Phe Leu Leu Leu
500 505 510
Asn Leu Lys Gln Thr Thr Val Lys Asn
515

<210> 11
<211> 251
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3342358CD1

<400> 11
Met Thr Asp Ser Ala Thr Ala Asn Gly Asp Asp Arg Asp Pro Glu
1 5 10 15
Ile Glu Leu Phe Val Lys Ala Gly Ile Asp Gly Glu Ser Ile Gly
20 25 30
Asn Cys Pro Phe Ser Gln Arg Leu Phe Met Ile Leu Trp Leu Lys
35 40 45

```
Gly Val Val Phe Asn Val Thr Thr Val Asp Leu Lys Arg Lys Pro
                50                  55                  60
Ala Asp Leu His Asn Leu Ala Pro Gly Thr His Pro Pro Phe Leu
                65                  70                  75
Thr Phe Asn Gly Asp Val Lys Thr Asp Val Asn Lys Ile Glu Glu
                80                  85                  90
Phe Leu Glu Glu Thr Leu Thr Pro Glu Lys Tyr Pro Lys Leu Ala
                95                  100                 105
Ala Lys His Arg Glu Ser Asn Thr Ala Gly Ile Asp Ile Phe Ser
                110                 115                 120
Lys Phe Ser Ala Tyr Ile Lys Asn Thr Lys Gln Gln Asn Asn Ala
                125                 130                 135
Ala Leu Glu Arg Gly Leu Thr Lys Ala Leu Lys Lys Leu Asp Asp
                140                 145                 150
Tyr Leu Asn Thr Pro Leu Pro Glu Glu Ile Asp Ala Asn Thr Cys
                155                 160                 165
Gly Glu Asp Lys Gly Ser Arg Arg Lys Phe Leu Asp Gly Asp Glu
                170                 175                 180
Leu Thr Leu Ala Asp Cys Asn Leu Leu Pro Lys Leu His Val Val
                185                 190                 195
Lys Ile Val Ala Lys Lys Tyr Arg Asn Tyr Asp Ile Pro Ala Glu
                200                 205                 210
Met Thr Gly Leu Trp Arg Tyr Leu Lys Asn Ala Tyr Ala Arg Asp
                215                 220                 225
Glu Phe Thr Asn Thr Cys Ala Ala Asp Ser Glu Ile Glu Leu Ala
                230                 235                 240
Tyr Ala Asp Val Ala Lys Arg Leu Ser Arg Ser
                245                 250


<210> 12
<211> 323
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1267774CD1

<400> 12
Met Gly Leu Phe Asp Arg Gly Val Gln Met Leu Leu Thr Thr Val
  1               5                   10                  15
Gly Ala Phe Ala Ala Phe Ser Leu Met Thr Ile Ala Val Gly Thr
                20                  25                  30
Asp Tyr Trp Leu Tyr Ser Arg Gly Val Cys Lys Thr Lys Ser Val
                35                  40                  45
Ser Glu Asn Glu Thr Ser Lys Lys Asn Glu Glu Val Met Thr His
                50                  55                  60
Ser Gly Leu Trp Arg Thr Cys Cys Leu Glu Gly Asn Ser Lys Gly
                65                  70                  75
Leu Cys Lys Gln Ile Asp His Phe Pro Glu Asp Ala Asp Tyr Glu
                80                  85                  90
Ala Asp Thr Ala Glu Tyr Phe Leu Arg Ala Val Arg Ala Ser Ser
                95                  100                 105
Ile Phe Pro Ile Leu Ser Val Ile Leu Leu Phe Met Gly Gly Leu
                110                 115                 120
```

Cys Ile Ala Ala Ser Glu Phe Tyr Lys Thr Arg His Asn Ile Ile
125 130 135
Leu Ser Ala Gly Ile Phe Phe Val Ser Ala Gly Leu Ser Asn Ile
140 145 150
Ile Gly Ile Ile Val Tyr Ile Ser Ala Asn Ala Gly Asp Pro Ser
155 160 165
Lys Ser Asp Ser Lys Lys Asn Ser Tyr Ser Tyr Gly Trp Ser Phe
170 175 180
Tyr Phe Gly Ala Leu Ser Phe Ile Ile Ala Glu Met Val Gly Val
185 190 195
Leu Ala Val His Met Phe Ile Asp Arg His Lys Gln Leu Arg Ala
200 205 210
Thr Ala Arg Ala Thr Asp Tyr Leu Gln Ala Ser Ala Ile Thr Arg
215 220 225
Ile Pro Ser Tyr Arg Tyr Arg Tyr Gln Arg Arg Ser Arg Ser Ser
230 235 240
Ser Arg Ser Thr Glu Pro Ser His Ser Arg Asp Ala Ser Pro Val
245 250 255
Gly Ile Lys Gly Phe Asn Thr Leu Pro Ser Thr Glu Ile Ser Met
260 265 270
Tyr Thr Leu Ser Arg Asp Pro Leu Lys Ala Ala Thr Thr Pro Thr
275 280 285
Ala Thr Tyr Asn Ser Asp Arg Asp Asn Ser Phe Leu Gln Val His
290 295 300
Asn Cys Ile Gln Lys Glu Asn Lys Asp Ser Leu His Ser Asn Thr
305 310 315
Ala Asn Arg Arg Thr Thr Pro Val
320

<210> 13
<211> 51
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1817329CD1

<400> 13
Met Asn Gln Gly Ser Gly Leu Asp Leu Leu Lys Ile Ser Tyr Gly
1 5 10 15
Lys Gly Ala Arg Arg Lys Asn Arg Phe Lys Gly Ser Asp Gly Ser
20 25 30
Thr Ser Ser Asp Thr Thr Ser Asn Ser Phe Val Arg Gln Val Arg
35 40 45
Val Leu Ser Ser Trp Phe
50

<210> 14
<211> 113
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 3273307CD1

<400> 14
Met Glu Gln Arg Lys Leu Asn Asp Gln Ala Asn Thr Leu Val Asp
1 5 10 15
Leu Ala Lys Thr Gln Asn Ile Met Tyr Asp Met Ile Ser Asp Leu
20 25 30
Asn Glu Arg Ser Glu Asp Phe Glu Lys Arg Ile Val Thr Leu Glu
35 40 45
Thr Lys Leu Glu Thr Leu Ile Gly Ser Ile His Ala Leu Pro Gly
50 55 60
Leu Ile Ser Gln Thr Ile Arg Gln Gln Gln Arg Asp Phe Ile Glu
65 70 75
Ala Gln Met Glu Ser Tyr Asp Lys His Val Thr Tyr Asn Ala Glu
80 85 90
Arg Ser Arg Ser Ser Ser Arg Arg Arg Arg Ser Ser Ser Thr Ala
95 100 105
Pro Pro Thr Ser Ser Glu Ser Ser
110

<210> 15
<211> 215
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3824833CD1

<400> 15
Met His Arg Asp Ala Trp Leu Pro Arg Pro Ala Phe Ser Leu Thr
1 5 10 15
Gly Leu Ser Leu Phe Phe Ser Leu Val Pro Pro Gly Arg Ser Met
20 25 30
Glu Val Thr Val Pro Ala Thr Leu Asn Val Leu Asn Gly Ser Asp
35 40 45
Ala Arg Leu Pro Cys Thr Phe Asn Ser Cys Tyr Thr Val Asn His
50 55 60
Lys Gln Phe Ser Leu Asn Trp Thr Tyr Gln Glu Cys Asn Asn Cys
65 70 75
Ser Glu Glu Met Phe Leu Gln Phe Arg Met Lys Ile Ile Asn Leu
80 85 90
Lys Leu Glu Arg Phe Gln Asp Arg Val Glu Phe Ser Gly Asn Pro
95 100 105
Ser Lys Tyr Asp Val Ser Val Met Leu Arg Asn Val Gln Pro Glu
110 115 120
Asp Glu Gly Ile Tyr Asn Cys Tyr Ile Met Asn Pro Pro Asp Arg
125 130 135
His Arg Gly His Gly Lys Ile His Leu Gln Val Leu Met Glu Glu
140 145 150
Pro Pro Glu Arg Asp Ser Thr Val Ala Val Ile Val Gly Ala Ser
155 160 165
Val Gly Gly Phe Leu Ala Val Val Ile Leu Val Leu Met Val Val
170 175 180

```
Lys Cys Val Arg Arg Lys Lys Glu Gln Lys Leu Ser Thr Asp Asp
                185                 190                 195
Leu Lys Thr Glu Glu Glu Gly Lys Thr Asp Gly Glu Gly Asn Pro
                200                 205                 210
Asp Asp Gly Ala Lys
                215


<210> 16
<211> 235
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2069907CD1

<400> 16
Met Phe Ile Trp Thr Ser Gly Arg Thr Ser Ser Ser Tyr Arg His
  1               5                  10                  15
Asp Glu Lys Arg Asn Ile Tyr Gln Lys Ile Arg Asp His Asp Leu
                 20                  25                  30
Leu Asp Lys Arg Lys Thr Val Thr Ala Leu Lys Ala Gly Glu Asp
                 35                  40                  45
Arg Ala Ile Leu Leu Gly Leu Ala Met Met Val Cys Ser Ile Met
                 50                  55                  60
Met Tyr Phe Leu Leu Gly Ile Thr Leu Leu Arg Ser Tyr Met Gln
                 65                  70                  75
Ser Val Trp Thr Glu Glu Ser Gln Cys Thr Leu Leu Asn Ala Ser
                 80                  85                  90
Ile Thr Glu Thr Phe Asn Cys Ser Phe Ser Cys Gly Pro Asp Cys
                 95                 100                 105
Trp Lys Leu Ser Gln Tyr Pro Cys Pro Gln Val Tyr Val Asn Leu
                110                 115                 120
Thr Ser Ser Gly Glu Lys Leu Leu Leu Tyr His Thr Glu Glu Thr
                125                 130                 135
Ile Lys Ile Asn Gln Lys Cys Ser Tyr Ile Pro Lys Cys Gly Lys
                140                 145                 150
Asn Phe Glu Glu Ser Met Ser Leu Val Asn Val Val Met Glu Asn
                155                 160                 165
Phe Arg Lys Tyr Gln His Phe Ser Cys Tyr Ser Asp Pro Glu Gly
                170                 175                 180
Asn Gln Lys Ser Val Ile Leu Thr Lys Leu Tyr Ser Ser Asn Val
                185                 190                 195
Leu Phe His Ser Leu Phe Trp Pro Thr Cys Met Met Ala Gly Gly
                200                 205                 210
Val Ala Ile Val Ala Met Val Lys Leu Thr Gln Tyr Leu Ser Leu
                215                 220                 225
Leu Cys Glu Arg Ile Gln Arg Ile Asn Arg
                230                 235


<210> 17
<211> 234
<212> PRT
<213> Homo sapiens
```

<220>
<221> misc_feature
<223> Incyte ID No: 2243917CD1

<400> 17
Met Ala Glu Asn His Cys Glu Leu Leu Ser Pro Ala Arg Gly Gly
1 5 10 15
Ile Gly Ala Gly Leu Gly Gly Gly Leu Cys Arg Arg Cys Ser Ala
20 25 30
Gly Leu Gly Ala Leu Ala Gln Arg Pro Gly Ser Val Ser Lys Trp
35 40 45
Val Arg Leu Asn Val Gly Gly Thr Tyr Phe Leu Thr Thr Arg Gln
50 55 60
Thr Leu Cys Arg Asp Pro Lys Ser Phe Leu Tyr Arg Leu Cys Gln
65 70 75
Ala Asp Pro Asp Leu Asp Ser Asp Lys Asp Glu Thr Gly Ala Tyr
80 85 90
Leu Ile Asp Arg Asp Pro Thr Tyr Phe Gly Pro Val Leu Asn Tyr
95 100 105
Leu Arg His Gly Lys Leu Val Ile Asn Lys Asp Leu Ala Glu Glu
110 115 120
Gly Val Leu Glu Glu Ala Glu Phe Tyr Asn Ile Thr Ser Leu Ile
125 130 135
Lys Leu Val Lys Asp Lys Ile Arg Glu Arg Asp Ser Lys Thr Ser
140 145 150
Gln Val Pro Val Lys His Val Tyr Arg Val Leu Gln Cys Gln Glu
155 160 165
Glu Glu Leu Thr Gln Met Val Ser Thr Met Ser Asp Gly Trp Lys
170 175 180
Phe Glu Gln Leu Val Ser Ile Gly Ser Ser Tyr Asn Tyr Gly Asn
185 190 195
Glu Asp Gln Ala Glu Phe Leu Cys Val Val Ser Lys Glu Leu His
200 205 210
Asn Thr Pro Tyr Gly Thr Ala Ser Glu Pro Ser Glu Lys Ala Lys
215 220 225
Ile Leu Gln Glu Arg Gly Ser Arg Met
230

<210> 18
<211> 301
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2597476CD1

<400> 18
Met Val Phe Thr Gln Ala Pro Ala Glu Ile Met Gly His Leu Arg
1 5 10 15
Ile Arg Ser Leu Leu Ala Arg Gln Cys Leu Ala Glu Phe Leu Gly
20 25 30
Val Phe Val Leu Met Leu Leu Thr Gln Gly Ala Val Ala Gln Ala
35 40 45
Val Thr Ser Gly Glu Thr Lys Gly Asn Phe Phe Thr Met Phe Leu

```
                50                  55                  60
Ala Gly Ser Leu Ala Val Thr Ile Ala Ile Tyr Val Gly Gly Asn
                65                  70                  75
Val Ser Gly Ala His Leu Asn Pro Ala Phe Ser Leu Ala Met Cys
                80                  85                  90
Ile Val Gly Arg Leu Pro Trp Val Lys Leu Pro Ile Tyr Ile Leu
                95                  100                 105
Val Gln Leu Leu Ser Ala Phe Cys Ala Ser Gly Ala Thr Tyr Val
                110                 115                 120
Leu Tyr His Asp Ala Leu Gln Asn Tyr Thr Gly Gly Asn Leu Thr
                125                 130                 135
Val Thr Gly Pro Lys Glu Thr Ala Ser Ile Phe Ala Thr Tyr Pro
                140                 145                 150
Ala Pro Tyr Leu Ser Leu Asn Asn Gly Phe Leu Asp Gln Val Leu
                155                 160                 165
Gly Thr Gly Met Leu Ile Val Gly Leu Leu Ala Ile Leu Asp Arg
                170                 175                 180
Arg Asn Lys Gly Val Pro Ala Gly Leu Glu Pro Val Val Val Gly
                185                 190                 195
Met Leu Ile Leu Ala Leu Gly Leu Ser Met Gly Ala Asn Cys Gly
                200                 205                 210
Ile Pro Leu Asn Pro Ala Arg Asp Leu Gly Pro Arg Leu Phe Thr
                215                 220                 225
Tyr Val Ala Gly Trp Gly Pro Glu Val Phe Ser Ala Gly Asn Gly
                230                 235                 240
Trp Trp Trp Val Pro Val Val Ala Pro Leu Val Gly Ala Thr Val
                245                 250                 255
Gly Thr Ala Thr Tyr Gln Leu Leu Val Ala Leu His His Pro Glu
                260                 265                 270
Gly Pro Glu Pro Ala Gln Asp Leu Val Ser Ala Gln His Lys Ala
                275                 280                 285
Ser Glu Leu Glu Thr Pro Ala Ser Ala Gln Met Leu Glu Cys Lys
                290                 295                 300
Leu
```

<210> 19
<211> 2994
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1568324CB1

<400> 19

```
gggaggcact gcccgtgttg ggatgcagaa gggatcagct tccagttgtc ttggagttga   60
tgactgacct ctacctctgc agggaaaggg gcgcagggag tctgaaagct ccaggagcag  120
cagaaggccc agtggccggt ctccaaccag tactgagaag cgcatgagct tcgagtccat  180
ttcttccctg ccagaggttg agccggaccc tgaggctggg agtgagcaag aggtattttc  240
tgctgtggaa gggcccagtg ccgaggagac gccttcagac acagaatctc cagaagtcct  300
ggagacacag cttgatgccc accagggcct tctggggatg gaccccccag gtgacatggt  360
ggacttcgtg gcagctgaga gcactgagga ccttaaggcc ctgagcagcg aggaggaaga  420
agaaatggga ggtgccgccc aggagcctga gagccttctg ccaccctctg tgctggacca  480
ggccagcgtc attgcggagc gatttgtcag cagcttctct cggcggagca gcgtggcaca  540
ggaggacagc aagtccagtg gctttgggag cccgcggctg gtcagccgga gcagcagcgt  600
```

gctcagcctg gagggcagcg agaagggcct ggcccggcat ggcagtgcca cagactccct 660
cagctgtcag ctctccccag aagtggacat cagtgtgggg gtggccacag aggacagccc 720
ttctgtcaat gggatggagc cccaagccc aggctgccca gtggagcctg accggtcttc 780
ctgcaagaag aaggaatcag cactctccac ccgagaccgg ctgttgctag acaagattaa 840
gagctattat gaaaatgcag aacaccatga tgcaggcttc agcgtccgtc gccgggagag 900
cctctcctac atccccaaag gactggtaag aaactccatc tccaggttca acagccttcc 960
ccggccagac ccagagccag tacctccagt ggggagcaag agacaggtgg gctcccggcc 1020
gacttcgtgg gccctgtttg agctcccagg accaagccag gcagtcaaag gggacccacc 1080
tcccatctca gatgctgagt tccgcccatc ttcagaaatt gtgaagatct gggagggaat 1140
ggagtcttcc ggagggagcc ctgggaaggg gccaggccag ggccaggcca atggctttga 1200
cctgcatgag ccactcttca tcctggagga gcatgagctg ggagccatca cagaggagtc 1260
ggccactgcc tccccggaaa gctcctctcc cactgagggg cgcagcccgg cccacctggc 1320
ccgggagctg aaagagctgg tgaaggagct gagcagcagt acccaggggg agctggtggc 1380
cccactgcac ccccgcatcg tgcagctctc ccacgtaatg gacagccacg tgagcgagcg 1440
cgtcaagaac aaggtctacc agctggcccg ccagtacagc ctccggatca agagcaacaa 1500
gccagtgatg gccaggccac cactgcagtg ggaaaaggtg gcccctgaga gggatgggaa 1560
gagccccact gtgccctgtc tacaggaaga ggctggagag ccattaggtg gcaaaggtaa 1620
gaggaagccg gtgctgtctc tatttgacta tgagcagctg atggcccagg agcacagccc 1680
tcccaagccc tcctcggctg ggagatgtc accacagcgt ttcttcttca acccgcctgc 1740
tgtcagccag aggaccacct cgcctggggg ccggccctcc gcccggagcc ccctcagccc 1800
cacagagacc ttcagctggc ccgacgtccg tgagctctgc tccaagtatg cctcccgcga 1860
tgaggcacgc cgagcagggg gcggccggcc ccgcggccca cccgtcaaca ggagccactc 1920
ggtgccggag aacatggtag agccacctct gtcgggcagg gtgggccgct gccgcagcct 1980
gagcaccaag aggggccggg gaggcggaga ggctgcccaa tcccctgggc ctctgcccca 2040
gagcaagccg gatggaggcg agaccctgta tgtcactgca gacctcaccc tggaggacaa 2100
ccggcgggtg attgtcatgg agaagggacc ccttcccagc cccactgcag ggctggagga 2160
gagcagtggc cagggaccaa gctcaccggt ggccctgctg ggcaggttc aggacttcca 2220
gcagtctgca gagtgccagc cgaaggaaga gggttccagg gacccggcag acccgagcca 2280
gcagggcaga gtgagaaacc ttagagagaa gttccaggcc ttgaactctg tcggttgatg 2340
ctgactcctg gggagggag gagtcatgtt ggaggttggg gaagaacctg ggcatccttc 2400
ccctcaagcc tgggctcatg gagcccctgc ccagggccct caggtgggcg gaaagtccat 2460
cccctccgcc cttcaggaag gatgctcccg tgtgcagggg tctcctgcct gtgccatcca 2520
ctggggctcg agacaatttc ccactcacct gtgaggccgg tgtggctgct tcccttgtaa 2580
atagttgttc tctggtaaga agccaaatat ttaagctcac ttcttcccag agagaggaag 2640
ctctgctcag gcctccagcg ttggctggcc atggccacag ccagatggag gagcccatcc 2700
ccaggagact caggcagtgg cctggagagg ctttgttctg taacggtgcc ttttcttagg 2760
gtccaggcag gaatgaagcc aataatttat tgctttccat tctgtggtat gatgtgcgtg 2820
tgcgtgagtg tgtggcccct gtttattccc ctcctgtcaa gaatgaagtg gattcagttc 2880
aggtactttt gagggttgtt gtgctgaccc tgtggttgtc gctgatgtac acacatttca 2940
ttatttgcca atggtgcaat aaccactgct gaccaaccca aaaaaaaaaa aaaa 2994

<210> 20
<211> 1298
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4094907CB1

<400> 20
cacgaccccc tgcctcttct tccagaagcc aatgggcaca gaagcaccaa ttctcccaca 60
atagtttcac ctgctattgt ttcccccacc caggacagtc ggcccaatat gtcaagacct 120
ctgatcacta gatcccctgc atctccactg aacaaccaag gcatccctac tccagcacaa 180
ctcacaaaat ccaatgcgcc tgtccacatt gatgtgggcg gccacatgta caccagcagc 240
ctggccaccc tcaccaaata ccctgaatcc agaatcggaa gactttttga tggtacagag 300

cccattgttt tggacagtct caaacagcac tatttcattg acagagatgg acagatgttc 360
agatatatct tgaattttct acgaacatcc aaactcctca ttcctgatga tttcaaggac 420
tacactttgt tatatgaaga ggcaaaatat tttcagcttc agcccatgtt gttggagatg 480
gaaagatgga agcaggacag agaaactggt cgattttcaa ggccctgtga gtgcctcgtc 540
gtgcgtgtgg ccccagacct cggagaaagg atcacgctaa gcggtgacaa atccttgata 600
gaagaagtat ttccagagat cggcgacgtg atgtgtaact ctgtcaatgc aggctggaat 660
cacgactcga cgcacgtcat caggtttcca ctaaatggct actgtcacct caactcagtc 720
caggtcctcg agaggttgca gcaaagagga tttgaaatcg tgggctcctg tgggggagga 780
gtagactcgt cccagttcag cgaatacgtc cttcggcggg aactgaggcg gacgccccgt 840
gtaccctccg tcatccggat aaagcaagag cctctggact aaatggacat atttcttatg 900
caaaaaggaa aacacacaca accaataact caaacaaaaa agggacattt atgtgcagtt 960
gggacagcaa accaagtcct ggacgtaaaa tcgaataaaa gacacattta tatccaatag 1020
agaccacacc tgtattcata tgggaacaat tggaatagtg atatcctcaa ggtgtaaaaa 1080
atatataaat atatatatat atgtcaaaag gtaggaaatg caaaaaagaa aaaaaaaaaa 1140
aggtgacagc cgcagttggt gctgtgatgg ccgtgaagtg tcctgggcct cccgaggcct 1200
ctgacaaata aacaagccat gagtggtgag gacacagtct ccttacagtt tccattgcca 1260
acaacagcca tccatatttc tttttccttt tgtctttc 1298

<210> 21
<211> 1877
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 518158CB1

<400> 21
gcctggccgt caccactccc cagagggcac agggctctgc tgtgcctcag agcaaaagtc 60
ccagagccag cagagcaggc tgacgacctg caagccacag tggctgccct gtgcgtgctg 120
cgaggtgggg gaccctgggc aggaagctgg ctgagcccca agaccccggg ggccatgggc 180
ggggatctgg tgcttggcct gggggccttg agacgccgaa agcgcttgct ggagcaggag 240
aagtctctgg ccggctgggc actggtgctg gcaggaactg gcattggact catggtgctg 300
catgcagaga tgctgtggtt cgggggggtgc tcgtgggcgc tctacctgtt cctggttaaa 360
tgcacgatca gcatttccac cttcttactc ctctgcctca tcgtggcctt tcatgccaaa 420
gaggtccagc tgttcatgac cgacaacggg ctgcgggact ggcgcgtggc gctgaccggg 480
cggcaggcgg cgcagatcgt gctggagctg gtggtgtgtg ggctgcaccc ggcgcccgtg 540
cggggcccgc cgtgcgtgca ggatttaggg gcgccgctga cctccccgca gccctggccg 600
ggattcctgg gccaagggga agcgctgctg tccctggcca tgctgctgct cggcctcacg 660
cttggcctct ggctgaccac cgcctgggtg ctgtccgtgg ccgagaggca ggctgttaat 720
gccactgggc acctttcaga cacactttgg ctgatcccca tcacattcct gaccatcggc 780
tatggtgacg tggtgccggg caccatgtgg ggcaagatcg tctgcctgtg cactggagtc 840
atgggtgtct gctgcacagc cctgctggtg gccgtggtgg cccggaagct ggagtttaac 900
aaggcagaga agcacgtgca caacttcatg atggatatcc agtataccaa agagatgaag 960
gagtccgctg cccgagtgct acaagaagcc tggatgttct acaaacatac tcgcaggaag 1020
gagtctcatg ctgcccgcag gcatcagcgc aagctgctgg ccgccatcaa cgcgttccgc 1080
caggtgcggc tgaaacaccg gaagctccgg gaacaagtga actccatggt ggacatctcc 1140
aagatgcaca tgatcctgta tgacctgcag cagaatctga gcagctcaca ccgggccctg 1200
gagaaacaga ttgacacgct ggcggggaag ctggatgccc tgactgagct gcttagcact 1260
gccctggggc cgaggcagct tccagaaccc agccagcagt ccaagtagct ggacccacga 1320
ggaggaacca ggctactttc cccagtactg aggtggtgga catcgtctct gccactcctg 1380
acccagccct gaacaaagca cctcaagtgc aaggaccaaa gggggccctg gcttggagtg 1440
ggttggcttg ctgatggctg ctggagggga cgctggctaa agtgggtagg ccttggccca 1500
cctgaggccc caggtgggaa catggtcacc cccactctgc ataccctcat caaaaacact 1560
ctcactatgc tgctatggac gacctccagc tctcagttac aagtgcaggc gactggaggc 1620
aggactcctg ggtccctggg aaagagggta ctaggggccc ggatccagga ttctgggagg 1680

cttcagttac cgctggccga gctgaagaac tgggtatgag gctggggcgg ggctggaggt 1740
ggcgccccct ggtgggacaa caaagaggac accatttttc cagagctgca gagagcacct 1800
ggtggggagg aagaagtgta actcaccagc ctctgctctt atctttgtaa taaatgttaa 1860
agccagaaaa aaaaaaa 1877

<210> 22
<211> 2517
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 602926CB1

<400> 22
gccaccccgc tctcctcgcc gccgcggcgg caggcgcggg gccgcgccgc gaggcctgat 60
ccctgcagcg cgggcaggcg gcgtcgcaga gcggagctca gctggatgcg tccggactcc 120
tgcaggtgag agtgattttc cagtgattgc tttggcctgt acaaccagag aacaggattc 180
ttcccttctt tttggccacc aaatgcctat gtgcaccaca cattccagtg tgctgagaag 240
ggcagagctt cttggatgat gatggacgtc ccaccgggca ggatgaaggc agagcgtgtg 300
gcatctccac ctcaagggtg cagcctgatc ttcctcttct cccttgccag ccagcactct 360
gccttctgta tccaccatgg tgtttggtga gtttttccat cgccctggac aagacgagga 420
acttgtcaac ctgaatgtgg ggggctttaa gcagtctgtt gaccaaagca ccctcctgcg 480
gtttcctcac accagactgg ggaagctgct tacttgccat tctgaagagg ccattctgga 540
gctgtgtgat gattacagtg tggccgataa ggaatactac tttgatcgga atccctcctc 600
gttcagatat gttttgaatt tttattacac ggggaagctg catgtcatgg aggagctgtg 660
cgtattctca ttctgccagg agatcgagta ctggggcatc aacgagctct tcattgattc 720
ttgctgcagc aatcgctacc aggaacgcaa ggaggaaaac cacgagaagg actgggacca 780
gaaaagccat gatgtgagta ccgactcctc gtttgaagag tcgtctctgt ttgagaaaga 840
gctggagaag tttgacacac tgcgatttgg tcagctccgg aagaaaatct ggattagaat 900
ggagaatcca gcgtactgcc tgtccgctaa gcttatcgct atctcctcct tgagcgtggt 960
gctggcctcc atcgtggcca tgtgcgttca cagcatgtcg gagttccaga atgaggatgg 1020
agaagtggat gatccggtgc tggaaggagt ggagatcgcg tgcattgcct ggttcaccgg 1080
ggagcttgcc gtccgcctgg ctgccgctcc ttgtcaaaag aaattctgga aaaaccctct 1140
gaacatcatt gactttgtct ctattattcc cttctatgcc acgttggctg tagacaccaa 1200
ggaggaagag agtgaggata ttgagaacat gggcaaggtg gtccagatcc tacggcttat 1260
gaggattttc cgaattctaa agcttgcccg gcactcggta ggacttcggt ctctaggtgc 1320
cacactgaga cacagctacc atgaagttgg gcttctgctt ctcttcctct ctgtgggcat 1380
ttccattttc tctgtgctta tctactccgt ggagaaagat gaccacacat ccagcctcac 1440
cagcatcccc atctgctggt ggtgggccac catcagcatg acaactgtgg gctatggaga 1500
caccccacccg gtcaccttgg cgggaaagct catcgccagc acatgcatca tctgtggcat 1560
cttggtggtg gcccttccca tcaccatcat cttcaacaag ttttccaagt actaccagaa 1620
gcaaaaggac attgatgtgg accagtgcag tgaggatgca ccagagaagt gtcatgagct 1680
accttacttt aacattaggg atatatatgc acagcggatg cacgccttca ttaccagtct 1740
ctcttctgta ggcattgtgg tgagcgatcc tgactccaca gatgcttcaa gcattgaaga 1800
caatgaggac atttgtaaca ccacctcctt ggagaattgc acagcaaaat gagcgggggt 1860
gtttgtgcct gtttctctta tcctttcccg acattaggtt aacacagctt tataaacctc 1920
agtgggttcg ttaaaatcat ttaattctca gggtgtacct ttcagccata gttggacatt 1980
cattgctgaa ttctgaaatg atagaattgt ctttattttt ctctgtgagg tcaattaaat 2040
gccttgttct gaaatttatt ttttacaaga gagagttgtg atatagtttg gaatataaga 2100
taaatggtat tgggtggggt ttgtggctac agcttatgca tcattctgtg tttgtcattt 2160
actcacattg agctaacttt aaattactga caagtagaat caaaggtgca gctgactgag 2220
acgacatgca tgtaagatcc acaaaatgag acaatgcatg taaatccatg ctcatgttct 2280
aaacatggaa actaggagcc taataaactt cctaattcag tatggtatac caaaaaatcc 2340
gggcggcctg cgactagcga gctcgtctga ccgggaatcc attccgcgac ggtacctgcc 2400
ggcgtaccag cctttcccat agtgagtcgg attagagctt ggcggaatca tggtcatagc 2460

cggttcccgt gggaaactgt tatccggcca caattccata caacactcga gccggga 2517

<210> 23
<211> 1154
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 922119CB1

<400> 23
ttaactcagc tgggagttga agagccgatg ggcagcaggc agacttgagt ctcctttctg 60
tccatgggct cgggccactg tctcaggtcc acccgtggct ccaaaatggt ctcctggtcc 120
gtgatagcaa agatccagga aatactgcag aggaagatgg tgcgagagtt cctggccgag 180
ttcatgagca catatgtcat gatggtattc ggccttggtt ccgtggccca tatggttcta 240
aataaaaaat atgggagcta ccttggtgtc aacttgggtt ttggcttcgg agtcaccatg 300
ggagtgcacg tggcaggccg catctctgga gcccacatga acgcagctgt gacctttgct 360
aactgtgcgc tgggccgcgt gccctggagg aagtttccgg tctatgtgct ggggcagttc 420
ctgggctcct tcctggcggc tgccaccatc tacagtctct tctacacggc cattctccac 480
ttttcgggtg gacagctgat ggtgaccggt cccgtcgcta cagctggcat ttttgccacc 540
taccttcctg atcacatgac attgtggcgg ggcttcctga atgaggcgtg gctgaccggg 600
atgctccagc tgtgtctctt cgccatcacg gaccaggaga acaacccagc actgccagga 660
acagaggcgc tggtgatagg catcctcgtg gtcatcatcg ggggtgtccct tggcatgaac 720
acaggatatg ccatcaaccc gtcccgggac ctgccccccc gcatcttcac cttcattgct 780
ggttggggca aacaggtctt cagcaatggg gagaactggt ggtgggtgcc agtggtggca 840
ccacttctgg gtgcctatct aggtggcatc atctacctgg tcttcattgg ctccaccatc 900
ccacgggagc ccctgaaatt ggaggattct gtggcgtatg aagaccacgg gataaccgta 960
ttgcccaaga tgggatctca tgaacccacg atctctcccc tcacccccgt ctctgtgagc 1020
cctgccaaca gatcttcagt ccaccctgcc ccacccttac atgaatccat ggccctagag 1080
cacttctaag cagagattat ttgtgatccc atccattccc caataaagca aggcttgtcc 1140
gacaaaaaaa aaaa 1154

<210> 24
<211> 1879
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2666782CB1

<400> 24
ggtaccgagg cccgagccgc gggagtcgag cgaaggcagc gccgaggccg cggtttcccc 60
ctgggcctcc ccagcagcag ccatgggcat caaattttta gaagttatca aaccattctg 120
tgcagttcta ccagaaattc agaaaccgga aaggaaaatc cagtttagag agaaggttct 180
gtggactgct ataacgctct tcatttcctt agtgtgttgt cagatcccac tgtttggaat 240
catgtcatca gattctgcag atcctttcta ctggatgaga gttattctgg cttccaatag 300
aggaacttta atggaattgg gtatctcccc aattgtaaca tctggtttga ttatgcagtt 360
gttagctgga gccaaaatca ttgaagttgg agatacaccg aaagatagag ctttattcaa 420
tggagcccag aaactgtttg gtatgatcat taccattggg caagccattg tgtatgtcat 480
gacggggatg tatggggacc ctgcagaaat gggtgccgga atctgtctcc tgatcatcat 540
tcagttgttt gttactagtt tgattgtgct actgttagat gagctactac agacaggtta 600
cagcttgggg tctgggattt ccctcgttat tgccaccaac atctgtgaga ccattgtctg 660
gaaggccttt agtcccacta ccattaacac tggcagaggt actgagtttg agggtgcagt 720
catagctctg ttccatttgt tggccaccag gacggacaaa gtccgagctt tacgggaggc 780

```
tttttatcgg cagaacttac ccaatctcat gaacctcatt gctacagttt ttgtgtttgc 840
tgttgttata tatttccaag gatttcgcgt tgatctgccc attaagtcgg cccgttaccg 900
aggacagtac agcagctacc ccatcaaact cttctacacc tccaacatcc ccatcatcct 960
ccagtcggcc ctggtgtcca acctgtatgt tatttcccag atgctgtctg ttcgatttag 1020
tggcaacttt ttagtaaatt tactaggaca gtgggccgat gtcagtgggg gaggacccgc 1080
acgttcttac ccagttggag gcctttgtta ctatctttct cctcctgagt ccatgggcgc 1140
catctttgag gatcctgtcc atgtcgttgt ttatatcatc ttcatgttgg ggtcatgtgc 1200
attcttctct aagacatgga ttgaagtgtc tggttcctca gccaaagatg tagctaaaca 1260
gctgaaagaa cagcagatgg taatgagggg ccaccgagat acctctatgg ttcatgagct 1320
taataggtac atccccaccg cagctgcgtt tggcggtttg tgcattggcg ccctgtcagt 1380
gctggctgac ttcctggggg ccattggatc tggcactgga attctgctag cagtcactat 1440
tatttaccag tattttgaaa tatttgttaa agaacaggcc gaagttggtg ggatgggtgc 1500
tttgtttttc taaatgttca aatatttcat tttgtgcgtg tgaaagggaa aacactttga 1560
cggatcgttt ttgtcagatg acactggtgg ctcccctttt ctcccctcac agtttcttgt 1620
ttcgagtgct gactgacccg tttctgaaat gggcaccgag ctaagtctgt gtgcagcatt 1680
agtacccgct gccttaaaac tcaagtttac attattcatt aaaaaaagta catctagtgt 1740
tgcctgtaat gctggaaacc agtgtatcta ccttgctgtg ttaaatcatg acagtgagac 1800
ggtgagatgg attcgttttg cacacaacat tcaaaacact tcatattgcc cccacttgtt 1860
gaaaaataaa tgtagttca                                             1879
```

<210> 25
<211> 1537
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2731369CB1

<400> 25

```
ctggcccagc cctggcagct gatgggagag cagatcgtcc aattcaagcc ttcctacggt   60
ctctttagaa gacatgggct gaaacttggg gtaatagagg ttggctggca tatccatgca 120
ggatgattgt cttcacatgt atctattacc ttgtagaata aggtagaccc tgatttttgga 180
acctgaagac caaagtgcaa gattagctct gctacttcca tctgtggacc attgggcagg 240
tatctctggg ccttcactta ctctttgtga aatgaggaca ggggcaatcc ctaccctacc 300
aagtcattgg gagtgaagac atgatgacac ggtgattgtg aaaagatttt gtcaatcgca 360
ccagcattaa gggtgcccat ctccaggttc cccaggcct caaggctccc aaggcctgag 420
tgggcaggta gcacccaggt atagaccttc cacgtgcagc acccaggaca cagccagcat 480
gaactgggca tttctgcagg gcctgctgag tggcgtgaac aagtactcca cagtgctgag 540
ccgcatctgg ctgtctgtgg tgttcatctt tcgtgtgctg gtgtacgtgg tggcagcgga 600
ggaagtgtgg gacgatgagc agaaggactt tgactgcaac accaagcagc ccggctgcac 660
caacgtctgc tacgacaact acttccccat ctccaacatc cgcctctggg ccctacagct 720
catcctggtc acgtgcccct cactgctcgt ggtcatgcac gtggcctacc gcgaggaacg 780
cgagcgcaag caccacctga aacacgggcc caatgccccg tccctgtacg acaacctgag 840
caagaagcgg ggcggactgt ggtggacgta cttgctgagc ctcatcttca aggccgccgt 900
ggatgctggc ttcctctata tcttccaccg cctctacaag gattatgaca tgccccgcgt 960
ggtggcctgc tccgtggagc cttgccccca cactgtggac tgttacatct cccggcccac 1020
ggagaagaag gtcttcacct acttcatggt gaccacagct gccatctgca tcctgctcaa 1080
cctcagtgaa gtcttctacc tggtgggcaa gaggtgcatg gagatcttcg gccccaggca 1140
ccggcggcct cggtgccggg aatgcctacc cgatacgtgc ccaccatatg tcctctccca 1200
gggagggcac cctgaggatg ggaactctgt cctaatgaag gctgggtcgg ccccagtgga 1260
tgcaggtggg tatccataac ctgcgagatc agcagataag atcaacaggt cccccccaca 1320
tgaggccacc caggaaaaaa ggcaggggca gtggcatcct tgccgtagca gggtggtgag 1380
gagggtggct gtgggggctc aggaagctcg cccaggggcc aatgtgggag gttgggggta 1440
gtttggtccc tgggtcctga gcctcagggg agggaggttg atagctactg gggattttgt 1500
atatggcaac agtatatgtc aaacctctta ataaatt                          1537
```

<210> 26
<211> 884
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> 790, 827, 860
<223> a or g or c or t, unknown, or other

<220>
<221> misc_feature
<223> Incyte ID No: 1375415CB1

<400> 26
gcgcctggag ccacacaggg atccggagcc tgggggaaaa gcggcgcggg agccggcacc 60
caccgctgga ggggcggcga cggcggccgt agcgacctcg ggaggcaagc ggagccgcca 120
tggccgagtt cccgtcgaaa gttagcacgc ggaccagcag tcctgcgcag ggcgccgaag 180
cctcggtgtc ggcgctgcgc ccggacctgg gcttcgtgcg ctcccgcctc ggggcgctca 240
tgctgctgca gctggtgctg gggctgctgg tgtgggcgct gattgcggac accccgtacc 300
acctgtatcc ggcctatggc tgggtgatgt tcgtcgctgt cttcctctgg ctggtgacaa 360
tcgtcctctt caacctctac ctgtttcagc tgcacatgaa gttgtacatg gttccctggc 420
cactggtgtt aatgatcttt aacatcagcg ccaccgttct ctacatcacc gccttcatcg 480
cctgctctgc ggcagttgac ctgacatccc tgaggggcac ccggccttat aaccagcgcg 540
cggctgcctc gttctttgcg tgtttggtga tgatcgccta tggagtgagt gccttcttca 600
gctaccaggc ctggcgagga gtaggcagca atgcggccac cagtcagatg gctggcggct 660
atgcctaaac cacctgtgcc acggccccct ctggggctga agccgccgct gggtcacaga 720
gcagggtcac cctgcaagcc tgaagctggg gagccctgcg tggagtcagc ccaacaggga 780
ctgcatttgn ctcctctctg cccgtcagac ataagctctc acagcgntaa ggaagcaggc 840
ccaggctggc agacatctcn gcttgcagga ggccaactgt gaga 884

<210> 27
<211> 3156
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2733282CB1

<400> 27
gttcaggaag aaaccatctg catccatatt gaaaacctga cacaatgtat gcagcaggct 60
cagtgtgagt gaactggagg cttctctaca acatgaccca aaggagcatt gcaggtccta 120
tttgcaacct gaagtttgtg actctcctgg ttgccttaag ttcagaactc ccattcctgg 180
gagctggagt acagcttcaa gacaatgggt ataatggatt gctcattgca attaatcctc 240
aggtacctga gaatcagaac ctcatctcaa acattaagga aatgataact gaagcttcat 300
tttacctatt taatgctacc aagagaagag tatttttcag aaatataaag attttaatac 360
ctgccacatg gaaagctaat aataacagca aaataaaaca agaatcatat gaaaaggcaa 420
atgtcatagt gactgactgg tatggggcac atggagatga tccatacacc ctacaataca 480
gagggtgtgg aaaagaggga aaatacattc atttcacacc taatttccta ctgaatgata 540
acttaacagc tggctacgga tcacgaggcc gagtgtttgt ccatgaatgg gcccacctcc 600
gttggggtgt gttcgatgag tataacaatg acaaaccttt ctacataaat gggcaaaatc 660
aaattaaagt gacaaggtgt tcatctgaca tcacaggcat ttttgtgtgt gaaaaaggtc 720
cttgccccca agaaaactgt attattagta agctttttaa agaaggatgc acctttatct 780
acaatagcac ccaaaatgca actgcatcaa taatgttcat gcaaagttat ctctgtggtg 840
aaatttgtaa tgccagtacc cacaaccaag aagcaccaaa cctacagaac cagatgtgca 900

gcctcagaag tgcatgggat gtaatcacag actctgctga ctttcaccac agctttccca 960
tgaacgggac tgagcttcca cctcctccca cattctcgct tgtagaggct ggtgacaaag 1020
tggtctgttt agtgctggat gtgtccagca agatggcaga ggctgacaga ctccttcaac 1080
tacaacaagc cgcagaattt tatttgatgc agattgttga aattcatacc ttcgtgggca 1140
ttgccagttt cgacagcaaa ggagagatca gagcccagct acaccaaatt aacagcaatg 1200
atgatcgaaa gttgctggtt tcatatctgc ccaccactgt atcagctaaa acagacatca 1260
gcatttgttc agggcttaag aaaggatttg aggtggttga aaaactgaat ggaaaagctt 1320
atggctctgt gatgatatta gtgaccagcg gagatgataa gcttcttggc aattgcttac 1380
ccactgtgct cagcagtggt tcaacaattc actccattgc cctgggttca tctgcagccc 1440
caaatctgga ggaattatca cgtcttacag gaggtttaaa gttctttgtt ccagatatat 1500
caaactccaa tagcatgatt gatgctttca gtagaatttc ctctggaact ggagacattt 1560
tccagcaaca tattcagctt gaaagtacag gtgaaaatgt caaacctcac catcaattga 1620
aaaacacagt gactgtggat aatactgtgg gcaacgacac tatgtttcta gttacgtggc 1680
aggccagtgg tcctcctgag attatattat ttgatcctga tggacgaaaa tactacacaa 1740
ataattttat caccaatcta actttcgga cagctagtct ttggattcca ggaacagcta 1800
agcctgggca ctggacttac accctgaaca atacccatca ttctctgcaa gccctgaaag 1860
tgacagtgac ctctcgcgcc tccaactcag ctgtgccccc agccactgtg gaagcctttg 1920
tggaaagaga cagcctccat ttcctcatc ctgtgatgat ttatgccaat gtgaaacagg 1980
gattttatcc cattcttaat gccactgtca ctgccacagt tgagccagag actggagatc 2040
ctgttacgct gagactcctt gatgatggag caggtgctga tgttataaaa aatgatggaa 2100
tttactcgag gtattttttc tcctttgctg caaatggtag atatagcttg aaagtgcatg 2160
tcaatcactc tcccagcata agcaccccag cccactctat tccagggagt catgctatgt 2220
atgtaccagg ttacacagca aacggtaata ttcagatgaa tgctccaagg aaatcagtag 2280
gcagaaatga ggaggagcga aagtggggct ttagccgagt cagctcagga ggctcctttt 2340
cagtgctggg agttccagct ggcccccacc ctgatgtgtt tccaccatgc aaaattattg 2400
acctggaagc tgtaaaagta gaagaggaat tgaccctatc ttggacagca cctggagaag 2460
actttgatca gggccaggct acaagctatg aaataagaat gagtaaaagt ctacagaata 2520
tccaagatga ctttaacaat gctattttag taaatacatc aaagcgaaat cctcagcaag 2580
ctggcatcag ggagatattt acgttctcac cccaaatttc cacgaatgga cctgaacatc 2640
agccaaatgg agaaacacat gaaagccaca gaatttatgt tgcaatacga gcaatggata 2700
ggaactcctt acagtctgct gtatctaaca ttgcccaggc gcctctgttt attcccccca 2760
attctgatcc tgtacctgcc agagattatc ttatattgaa aggagtttta acagcaatgg 2820
gtttgatagg aatcatttgc cttattatag ttgtgacaca tcatacttta agcaggaaaa 2880
agagagcaga caagaaagag aatggaacaa aattattata aataaatatc caaagtgtct 2940
tccttcttag atataagacc catggccttc gactacaaaa acatactaac aaagtcaaat 3000
taacatcaaa actgtattaa aatgcattga gtttttgtac aatacagata agatttttac 3060
atggtagatc aacaaattct ttttgggggt agattagaaa acccttacac tttggctatg 3120
aacaaataat aaaaattatt ctttaaaaaa aaaaaa 3156

<210> 28
<211> 1774
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3148427CB1

<400> 28
cgctgctata ccgtgcaccc gcgtgctcgt gagggaactc gggctcgctt tggtggggag 60
aaaaatccat gcgctaatac tcggttccca gcttctgcaa aagaaataca aagagtatga 120
gaaagacgtt gccatagaag aacatcgatg gcctccaact ggtaaagaag ctggcaaaga 180
acatggaaga gatgtttcac aagaagtctg aggccgtcag gcgtctggtg gaggctgcag 240
aagaagcaca cctgaaacat gaatttgatg cagacttaca gtatgaatac ttcaatgctg 300
tgctgataaa tgaaagggac aaagacggga attttttgga gctgggaaag gaattcatct 360
tagccccaaa tgaccatttt aataatttgc ctgtgaacat cagtctaagt gacgtccaag 420

taccaacgaa catgtacaac aaagaccctg caattgtcaa tggggtttat tggtctgaat 480
ctctaaacaa agtttttgta gataactttg accgtgaccc atctctcata tggcagtact 540
ttggaagtgc aaagggcttt tttaggcagt atccggggat taaatgggaa ccagatgaga 600
atggagtcat tgccttcgac tgcaggaacc gaaaatggta catccaggca gcaacttctc 660
cgaaagacgt ggtcatttta gttgacgtca gtggcagcat gaaaggactc cgtctgacta 720
tcgcgaagca aacagtctca tccattttgg atacacttgg ggatgatgac ttcttcaaca 780
taattgctta taatgaggag cttcactatg tggaaccttg cctgaatgga actttggtgc 840
aagccgacag gacaaacaaa gagcacttca gggagcatct ggacaaactt ttcgccaaag 900
gaattggaat gttggatata gctctgaatg aggccttcaa cattctgagt gatttcaacc 960
acacgggaca aggaagtatc tgcagtcagg ccatcatgct cataactgat ggggcggtgg 1020
acacctatga tacaatcttt gcaaaataca attggccaga tcgaaaggtt cgcatcttca 1080
catacctcat tggacgagag gctgcgtttg cagacaatct aaagtggatg gcctgtgcca 1140
acaaaggatt ttttacccaa atctccacct tggctgatgt gcaggagaat gtcatggaat 1200
accttcacgt gcttagccgg cccaaagtca tcgaccagga gcatgatgtg gtgtggaccg 1260
aagcttacat tgacagcact ctccctcagg cacaaaagct gactgatgat cagggccccg 1320
tcctgatgac cactgtagcc atgcctgtgt ttagtaagca gaacgaaacc agatcgaagg 1380
gcattcttct gggagtggtt ggcacagatg tcccagtgaa agaacttctg aagaccatcc 1440
ccaaatacaa gttagggatt cacggttatg cctttgcaat cacaaataat ggatatatcc 1500
tgacgcatcc ggaactcagg ctgctgtacg aagaaggaaa aaagcgaagg aaacctaact 1560
atagtagcgt tgacctctct gaggtggagt gggaagaccg agatgacgtg ttgagaaatg 1620
ctatggtgaa tcgaaagacg gggaagtttt ccatggaggt gaagaagaca gtggacaaag 1680
ggtacatttt ttctcaaaca tttttgctgc ttaatttaaa acaaaccact gtgaaaaatt 1740
agctttgaaa gctatatctg gaataaatga cttc 1774

<210> 29
<211> 1505
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3342358CB1

<400> 29
cgcggatcct gtgacacctc cgggcagccc ggcacttgtt gctcccacga cctgttgtca 60
ttcccttaac ccggctttcc ccgtggcccc ccgcctcctc ccggcttcgc tccttttcat 120
gtgagcatct gggacactga tctctcagac cccgctgctc gggctggaga atagatggtt 180
ttgtgaaaaa ttaaacaccg ccctgaagag gagccccgct gggcagcggc aggagcgcag 240
agtgctggcc caggtgctgc agaggtggcg cctccccggc ccgggacggt agccccgggc 300
gccaacggca tgacagactc ggcgacagct aacggggacg acagggaccc cgagatcgag 360
ctctttgtga aggctggaat cgatggagaa agcatcggca actgtccttt ctctcagcgc 420
ctcttcatga tcctctggct gaaaggagtc gtgttcaatg tcaccactgt ggatctgaaa 480
agaaagccag ctgacctgca caacctagcc cccggcacgc acccgccctt cctgaccttc 540
aacggggacg tgaagacaga cgtcaataag atcgaggagt tcctggagga gaccttgacc 600
cctgaaaagt accccaaact ggctgcaaaa caccgggaat ccaacacagc gggcatcgac 660
atcttttcca agttttctgc ctacatcaaa aataccaagc agcagaacaa tgctgctctt 720
gaaagaggcc taaccaaggc tctaaagaaa ttggatgact acctgaacac ccctctacca 780
gaggagattg acgccaacac ttgtggggaa gacaaggggt cccggcgcaa gttcctggat 840
ggggatgagc tgaccctggc tgactgcaat ctgttgccca agctccatgt ggtcaagatt 900
gtggccaaga aataccgcaa ctatgatatc ccggctgaga tgacaggcct gtggcggtac 960
ctcaagaacg cctatgcccg tgatgagttc accaacacct gtgcagctga cagtgagatc 1020
gagttggcct acgctgatgt cgccaaacgc ctcagccgat cctgagcaca gccattttgc 1080
cccatccccg ctgcagaagg actcaaccac tcccctaaga ctccagcttc atagactcct 1140
ctgtatcact gccttgaggc gcacttttta taatcaagcc tcatcttgct ggtatcatgg 1200
gaactccagc ctgctatctt tcatgaaggt cagcaccatc ctgggcctcc tcacataggg 1260
atctagcaga aatgatagac acagtccacc tttcggccgg ccagcctgat ctggggctca 1320

gcatgtttgg gggtcagtca gtgttgggag agcccacata tggggattgc cattaggttt 1380
tttttgccat tatcaaaaat tacttttcaa gaggctttag ggaacatggc aacaacactt 1440
ctttttctaa cctccttttc ggcctatacc acaaggggca ggggcaaacg gcattttcat 1500
tcttt 1505

<210> 30
<211> 1478
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1267774CB1

<400> 30
gaaactggag accagaattt tagaaaaaga gattaaggca tctcacttgg tggggtgggg 60
ggtgtctttt tatttttttt ttccttttct ttttaaaaaa aaaacactgc aactggaaca 120
gtttctgatc tcaaaaggca agcctcttcc cgtgtgatct ttataattta cactcttttc 180
cgtgagcttt cttacctccc tttttttata actctccata ttctctattc atacatatat 240
ccattatatt agtagtggaa taatttttat ttttatttat tttttttggc tttagtactt 300
gcaccctcac acacactctc ccgagaacca gaagtcggtt gggtgtttat ataatgaaga 360
attatggggc tgtttgatcg aggtgttcaa atgcttttaa ccaccgttgg tgctttcgct 420
gccttcagcc tgatgaccat agctgtggga accgactatt ggctctactc cagaggggtt 480
tgcaagacca aaagtgtcag tgagaatgaa accagcaaaa agaacgagga agttatgacc 540
cattccggat tatggagaac ctgctgccta gaagggaatt caaaaggtct gtgcaagcaa 600
attgatcact tcccagagga tgcagattac gaagctgaca cagcagaata tttcctccgg 660
gccgtgaggg cctccagcat ttcccaatc ctgagtgtga ttctgctttt catgggtggc 720
ctctgcatcg cagccagcga gttctacaaa actcgacaca acatcatcct gagtgccggc 780
atcttcttcg tgtctgcagg tctgagtaac atcattggca tcatagtgta catatctgcc 840
aatgccggag acccctccaa gagcgactcc aaaaagaata gttactcata cggctggtcc 900
ttctacttcg ggccctgtc cttcatcatc gccgagatgg tcggggtgct ggcggtgcac 960
atgtttatcg accggcacaa acagctgcgg gccacggccc gcgccacgga ctacctccag 1020
gcctctgcca tcacccgcat ccccagctac cgctaccgct accagcgccg cagccgctcc 1080
agctcgcgct ccacggagcc ctcacactcc agggacgcct cccccgtggg catcaagggc 1140
ttcaacaccc tgccgtccac ggagatctcc atgtacacgc tcagcaggga ccccctgaag 1200
gccgccacca cgcccaccgc cacctacaac tccgacaggg ataacagctt cctccaggtt 1260
cacaactgta tccagaagga gaacaaggac tctctccact ccaacacagc caaccgccgg 1320
accacccccg tataaagacc gcgggcctcg ccagaagacc gcgggaggag ggcgcggtcc 1380
ccgggggcgg ggcggggcgg ggagacccag accctccgct gggagacctt ccaaaagcaa 1440
aaacaaaaaa caaaaaaaac aagtatacag gagagaga 1478

<210> 31
<211> 1971
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1817329CB1

<400> 31
cgggcggggg agacctgcca gtcctcccag acttctcccg ggttgctcca gctggccctc 60
ctcgcccctt cccgggagag gcacatggag agacatgaat caggggagtg gactggacct 120
gctgaagatc tcatatggaa aaggagccag aaggaaaaac agatttaaag gatctgatgg 180
aagcacgtca tctgatacta cctcaaatag ttttgttcgc caggtaagag ttttaagttc 240
atggttttga taagtacctt aaaatgactt tagattttta aaggtgggtt tcactctttt 300

ccttaagatt tatgtagtat gattgtactt acttttaatt gaagtgaaga cggtgctgtt 360
taatgtcata attaaaatat tttaatctta agtgagacta agtaataatt acactctttc 420
ctggcaagtt gaggaaagag aagtgtggca ttcattacag aggatttctt tgaaacagtg 480
agagaactcc agcagaaatg attatggatt tggggggatt gttttttttt ttttttttca 540
gtgggacaga gagctgatgt gcatctgtat ccctacctgt gagatactgg ggttcttcta 600
gttgagcttt cttcttctca cttgggcttg cattgaaaac tagaaaatca ttcttggtct 660
tgaaatggtt gaggctatcg aggttaaaca gaacatgcag tatcactgag acataattta 720
aaccatttct gcctttggag acatagcctt cgtctcattt aatgtgtagt atcttcctgc 780
caagcagcct ggaacagtcc cctaaacaga cagtgccagg ctccctaaca tataaacacg 840
cttatatatg gtaacagaaa ggaattccat taccagagga agagcataga ggcagctgga 900
cctctgaggg ttgtggttat gtgttgagaa tgaagtctca tcaggaatga gtgagcttgt 960
ttcttcctct aacccctctc catggggtgg aaagtagggg cagagcatgc agcagagaat 1020
ctgttcttgt ggcccaggga tgtccagtgt ctgatcagta ctctcatcat ctcttgtcaa 1080
tgacagctca ctctacggaa cagtggtagt caactggaaa agctgcctcg tgtcatttgg 1140
atgccacttt ccccagtgcc gcttgagtcc atcatcatgg ccatagattc cacccctgcc 1200
ccctgtttgt gtggcaggac tgtttcctat aaatcctata tgaccttttg ggttttattt 1260
ttattgggga cagttacatt tccctagctg tctaccctta ttggctcctt gtggcactcc 1320
ccaagtgcct ccctcatgtt tccttcccac agttagctgc agttgaactt gaacttgtcc 1380
acctgcagca tcggtgggga ttttgatctt ggctggttgc tgcttctttg cactgtcctt 1440
agacgaagga tgatctgtcc ccagccatca gtccccgctt gcaacatttg agtatgccag 1500
tggtacttcc aagtaacttg gcaactggaa aaaaactggt cctggtccgt gccaattggg 1560
aattgttgtg gttgcaggaa gtgagaaaaa gagttactta agaaggggca aagctttttt 1620
ggttagaaac atttagaaag aaaaaagatg aagccacgtg aggctcaacc ctagagctat 1680
gttctaatgg tgtcagcagg agctaggaag gctgtcagag gagtggaagg tgtatgcgtt 1740
ggcctcatgt tttcgtgggc agaatccagg catcatggtc ctgctgatga agggagggcc 1800
ttgggacata gtgacttggg agagtcagtt aaaggaaagt tacccatcgc cttttccatt 1860
tcaagctatt tattctgcct ctgagacaaa gagattgaac ctggagagct agagttagaa 1920
tctacaactc tgagtcatca tgcaggacct gaaacagacc aaaaaaaaaa a 1971

<210> 32
<211> 1424
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3273307CB1

<400> 32
cccttctcat ttgtcccatc catttgaaat acagtgtact taaatttcca acaaaatcat 60
aacgtagact ctgtcagtta attaaagtat actaacaagc atacaaattg ctaagtcatt 120
ctgggaggct atcagtgctg gatgtactta tacattcagc actggaaaca taccсttcca 180
gtagtgcaca gccagaaata ccaagcaggc ttgcaaagga caagattaac cttggcctca 240
tggtattttg tctttctgca gggaaaggaa gatcatggag acaacaattg caatacacgg 300
tggagaactc aacacccctt gattaaagaa aaaaaatagt tgaataaaaa agaagtttcc 360
tgaaggagta agatagaatt taccacagaa agaaagtggt agtatctccc tatttcaatt 420
taagtgttgt tcccaggcag gtacaagtac ttctaaaaag gcatagatta aagaaaaatg 480
gtatttggag tcatggttac taaatgaagc tacagctcac catgatcctt gggatgaaga 540
ctatgacccc cagcaaagtt acaaaggatt ctgctctgga atttatcaac tgctttgttt 600
gttctcttaa caggtaaaaa atgcagctgc caatgtactc agggaaacat ggctaattta 660
caaaaataca aagctagtga aaaagataga tcatgcaaaa gtaagaaaac atcaacgaaa 720
attcctgcaa gctattcatc aattaagaag tgtaaaaatg gagcagagga aactgaatga 780
ccaagcaaac actttggtgg acttggcaaa gacccagaac atcatgtatg atatgatttc 840
tgacttaaac gaaaggagtg aagacttcga gaagaggatt gttaccctgg aaacaaaact 900
agagactttg attggtagca tccacgccct ccctgggctc ataagccaga ccatcaggca 960
gcagcagaga gatttcattg aggctcagat ggagagctac gacaagcacg tcacttacaa 1020

tgctgagcgg tcccggtcct cgtccaggag gcggcggtcc tcttccacag caccaccaac 1080
ttcatcagag agtagctaga agagaataag ttaaccacaa aataagactt tttgccatca 1140
tatggtcaat attttagctt ttattgtaaa gcccctatgg ttctaatcag cgttatccgg 1200
gttctgatgt cagaatcctg ggaacctgaa cactaagttt taggccaaaa tgagtgaaaa 1260
ctcttttttt ttctttcaga tgcacaggga atgcacctat tattgctata tagattgttc 1320
ctcctgtaat ttcactaact ttttattcat gcacttcaaa caaactttac tactacatta 1380
tatgatatat aataaaaaaa gttaatttct gcacaaaaaa aaaa 1424

<210> 33
<211> 1224
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3824833CB1

<400> 33
cggccatatt agagagatgg aaataaagct tccttaatgt tgtatatgtc tttgaagtac 60
atccgtgcat tttttttag catccaacca ttcctccctt gtagttctcg cccccctcaaa 120
tcaccctctc ccgtagccca cccgactaac atctcagtct ctgaaaatgc acagagatgc 180
ctggctacct cgccctgcct tcagcctcac ggggctcagt ctctttttct ctttggtgcc 240
accaggacgg agcatggagg tcacagtacc tgccaccctc aacgtcctca atggctctga 300
cgcccgcctg ccctgcacct tcaactcctg ctacacagtg aaccacaaac agttctccct 360
gaactggact taccaggagt gcaacaactg ctctgaggag atgttcctcc agttccgcat 420
gaagatcatt aacctgaagc tggagcggtt tcaagaccgc gtggagttct cagggaaccc 480
cagcaagtac gatgtgtcgg tgatgctgag aaacgtgcag ccggaggatg aggggattta 540
caactgctac atcatgaacc cccctgaccg ccaccgtggt catggcaaga tccatctgca 600
ggtcctcatg gaagagcccc ctgagcggga ctccacggtg gccgtgattg tgggtgcctc 660
cgtcgggggc ttcctggctg tggtcatctt ggtgctgatg gtggtcaagt gtgtgaggag 720
aaaaaaagag cagaagctga gcacagatga cctgaagacc gaggaggagg gcaagacgga 780
cggtgaaggc aacccggatg atggcgccaa gtagtgggtg gccggcctgc agcctcctct 840
aggggttgca cccagcgctc cctcaggagg gccttggcct ggcacggctg tgctcctccc 900
ctgctcccag cccagagcag ccatcaggct ggaggtgacg atgagttcct gaaacttgga 960
ggggcatgtt aaagggatga ctgtgcattc cagggcactg acggaaagcc agggctgcag 1020
gcaaagctgg acatgtgccc tggcccagga ggccatgttg ggccctcgtt tccattgcta 1080
gtggcctcct tggggctcct gttggctcct aatcccttag gactgtggat gaggccagac 1140
tggaagagca gctccaggta ggggggccatg tttcccagcg ggacccacc aacagaggcc 1200
agtttcaaag tcagctgagg ggct 1224

<210> 34
<211> 1300
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2069907CB1

<400> 34
gctgtttacc ttgcatggtt gactgctctc ttctcacatt gtgtgccagg aagggcactg 60
caccttggta taaatgctgc tgggcaccgt tctgttttct ttcttttctt aatcctatcc 120
aagtatgcag tacgctcttg ggtcgtctca tgagacccag gggcatgttg gaaagaactg 180
agagaaagag caacaaagcg gcgagtggtg tgagagggca gcacgcgctg tggggcccctt 240
ccagagaaat gtactgaaaa agtctacgca atgtctggga tttgctaaac aatacctgga 300
aagcagacag gtctttttgc cattcctcca ggacatccac cataaggaaa ggagaccctg 360

gaccaacatt ctctaagatg tttatatgga ccagtggccg gacctcttca tcttatagac 420
atgatgaaaa aagaaatatt taccagaaaa tcagggacca tgacctcctg gacaaaagga 480
aaacagtcac agcactgaag gcaggagagg accgagctat tctcctggga ctggctatga 540
tggtgtgctc catcatgatg tattttctgc tgggaatcac actcctgcgc tcatacatgc 600
agagcgtgtg gaccgaagag tctcaatgca ccttgctgaa tgcgtccatc acggaaacat 660
ttaactgctc cttcagctgt ggtccagact gctggaaact ttctcagtac ccctgccccc 720
aggtgtacgt taacctgact tcttccgggg aaaagctcct cctctaccac acagaagaga 780
caataaaaat caatcagaag tgctcctata tacctaaatg tggaaaaaat tttgaagaat 840
ccatgtccct ggtgaatgtt gtcatggaaa acttcaggaa gtatcaacac ttctcctgct 900
attctgaccc agaaggaaac cagaagagtg ttatcctaac caaactctac agttccaacg 960
tgctgttcca ttcactcttc tggccaacct gtatgatggc tgggggtgtg gcaattgttg 1020
ccatggtgaa acttacacag tacctctccc tactatgtga gaggatccaa cggatcaata 1080
gataaatgca aaaatggata aaataatttt tgttaaagct caaatactgt ttctttcat 1140
tcttcaccaa agaaccttaa gtttgtaacg tgcagtctgt tatgagttcc ctaatatatt 1200
cttatatgta gagcaataat gcaaaagctg ttctatatgc aaacatgatg tctttattat 1260
tcaggagaat aaataactgt tttgtgttaa aaaaaaaaa 1300

<210> 35
<211> 1060
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2243917CB1

<400> 35
gcttgctggg atcatggcgg agaatcactg cgagctcctg tcgccggccc ggggcggcat 60
cggggcgggg ctgggggcg gcctgtgccg ccgctgcagc gctgggctcg gcgccctggc 120
ccagcgccct ggcagcgtgt ccaagtgggt ccgactcaac gtcggcggca cctacttcct 180
caccactcgg cagaccctgt gccgggaccc gaaatccttc ctgtaccgct tatgccaggc 240
cgatcccgac ctggactcag acaaggatga aacaggcgcc tatttaatcg acagagaccc 300
cacctacttt gggcctgtgc tgaactacct gagacacggc aagctggtga ttaacaaaga 360
cctcgcggag gaaggagtgt tggaggaagc agaattttac aatatcacct cattaataaa 420
acttgtaaag gacaaaatta gagaacgaga cagcaaaaca tcgcaggtgc ctgtgaagca 480
tgtgtaccgt gtgctgcagt gccaggagga ggagctcacg cagatggtgt ccaccatgtc 540
cgacggctgg aagttcgagc agttggtcag catcggctcc tcttacaact atggaacga 600
agaccaagcc gagttcctct gtgtggtgtc caaggagctg cacaacaccc cgtacggtac 660
ggccagcgag cccagcgaga aggccaagat tttgcaagaa cgaggctcaa ggatgtgagg 720
gacacagtat tgacagctga agaaatgatt tacgttttcc cgagatgtaa tgaactgcca 780
tgtccaggaa gcttggctgt gagaagaaac ctgcttttga tcatttttct agagatctgg 840
gtgtgaatcc ttttttgcct ctgaggtggg tggtgagaga cgggcccagc tgtccaaggc 900
cagacgtccc caagttgggg gagcacggcg gccgggtggg cgctgcctct tgggggggcc 960
tcgctctgtt ttttccaagt gccacgtggg actgaggcag acactcccag tcagcccgct 1020
cgatcctgaa gatcgtgtga aggaagcgtt cttggtgcta 1060

<210> 36
<211> 1815
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2597476CB1

<400> 36

```
gtgcctatgc agacagaggg agcagtgaat agcaataggg tgtttccacc atggtcttca   60
ctcaggcccc ggctgaaatc atgggccacc tccggatacg cagcctcctg gcccggcagt  120
gcctggcaga gtttctgggt gtgtttgtac tcatgctcct cacccaagga gctgtggccc  180
aggctgtcac cagtggagaa accaaaggca acttcttcac catgtttctg gctggctctc  240
tggccgttac gatagccatc tacgtgggtg gtaacgtctc aggggcccac ctgaatccag  300
ccttctccct ggccatgtgc atcgttggac gcctcccctg ggtcaagctc cccatttaca  360
tcttggtgca gttgctgtct gctttctgtg cttcgggagc cacctatgtt ctctaccatg  420
atgccctaca gaactataca ggtgggaacc tgacagtgac tggccccaag gagacagcct  480
ccatttttgc cacctatcct gcccccctatc tgtccctgaa caatggcttc ctggatcagg  540
ttctgggcac tgggatgctg attgtggggc tcttggccat cctggacaga cggaacaagg  600
gagtccctgc gggtctggag cctgtggtgg tggggatgct gatcctggcc ctcgggttat  660
ccatgggtgc caactgcggg attccactca accctgcccg ggacctgggc ccacgtctct  720
tcacctacgt ggctggctgg ggtcctgaag tcttcagtgc tggtaatggc tggtggtggg  780
tgcctgtggt ggcccctctg gtgggggcca ccgttggcac agccacttac cagctgttgg  840
tggctctgca ccaccctgag ggcccagagc cagctcagga tctggtgtct gctcaacaca  900
aagcctcaga gttggaaact cctgcctcag ctcagatgct ggagtgtaag ctatgattag  960
gacaaccctc acttcactca tggaccctgg agccagccac tgaccccgcc tgggaacaac 1020
agtcattctt cctctttgtt aatgtgccag aacctgggag gcttctctgt ttatctgttt 1080
ggcatccctt cctcctaaac taagaaggat cctggacagg gagaagtgga ggaggataag 1140
gtaccaggac tcaggcttct catcccctcc tcccgcaaag cggttttctg accctcaggg 1200
cctctcggaa tgtagttgct cgaggtaacc gctagagggt gcgcacctgg atgctggatg 1260
gggacggctg cgggcatctg cagggtggag gggccacca tccagtgtag ggcacaaccc 1320
tggggactgc cctccatagc ctgtcccgac tgccgactcc tagctctcat cgcctcggcg 1380
cctcccacct tcaccctctc ggggatgcct ccccaagagg gtagttaggg gtggggaagc 1440
cgcctccacc caggggggcgt ggtgggggcg gagggaagga gggcggcggg gcacagagac 1500
agagagcaag gctgtgaaac tgaggcaccg ttcctagaca tctcggtgct gtgtcgttca 1560
ttcaaggaga gttgagatac agtgaaatga gccagggcga ggagggaggg tgaaggaacg 1620
gagggcgggc ggctccgagg agcgagagtc gggctgaggg caacctggcg ccagggaaaa 1680
ttctggttat tcaccacttc tacagctctc ctgccgctcc ctgcagagga tgctcgtttt 1740
gcagagaagg cagtgttcct ctattccctt cttccgaatt aaaaataccc cctcagagcg 1800
aaaaaaaaaa aaaaa                                                  1815
```

<210> 37
<211> 315
<212> PRT
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> GenBank ID No: g2924369

<400> 37

```
Met Leu Gly Trp Val Gln Arg Val Leu Pro Gln Pro Pro Gly Thr
  1               5                  10                  15
Pro Gln Lys Thr Glu Glu Gly Ala Gly Pro Gln Pro Glu Thr Glu
                 20                  25                  30
Ser Lys Pro Glu Ala Asn Pro Gln Pro Glu Pro Glu Val Gln Pro
                 35                  40                  45
Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Ala
                 50                  55                  60
Pro Glu Glu Ala Ala Pro Glu Val Gln Thr Leu Pro Pro Glu Glu
                 65                  70                  75
Pro Val Glu Gly Glu Asp Val Ala Glu Ala Gly Pro Ser Leu Gln
                 80                  85                  90
Glu Thr Gln Glu Ala Asp Pro Pro Gln Pro Thr Ser Gln Ala Gln
                 95                 100                 105
```

Val Ala Val Val Lys Val Asn Arg Pro Ser Ser Trp Met Leu Ser
110 115 120
Trp Phe Trp Lys Gly Met Glu Lys Val Val Pro Gln Pro Val Tyr
125 130 135
Ser Ser Ser Gly Gly Gln Asn Leu Ala Ala Gly Glu Gly Gly Pro
140 145 150
Asp Gln Asp Gly Ala Gln Thr Leu Glu Pro Cys Gly Thr Gly Asp
155 160 165
Pro Gly Ser Glu Asp Gly Ser Asp Lys Thr Ser Lys Thr Gln Asp
170 175 180
Thr Glu Pro Ser Leu Trp Leu Leu Arg Trp Leu Glu Leu Asn Leu
185 190 195
Glu Lys Val Leu Pro Gln Pro Pro Thr Pro Ser Gln Ala Trp Lys
200 205 210
Val Glu Pro Glu Gly Ala Val Leu Glu Pro Asp Pro Pro Gly Thr
215 220 225
Pro Met Glu Val Glu Pro Thr Glu Asn Pro Ser Gln Pro Asn Pro
230 235 240
Gly Pro Val Glu Pro Glu Glu Glu Pro Ala Ala Glu Pro Gln Pro
245 250 255
Gly Phe Gln Ala Ser Ser Leu Pro Pro Pro Gly Asp Pro Val Arg
260 265 270
Leu Ile Glu Trp Leu Leu His Arg Leu Glu Met Ala Leu Pro Gln
275 280 285
Pro Val Leu His Gly Lys Ala Ala Glu Gln Glu Pro Ser Cys Pro
290 295 300
Gly Thr Cys Asp Val Gln Thr Arg Ala Thr Ala Ala Gly Gly Leu
305 310 315

<210> 38
<211> 490
<212> PRT
<213> Drosophila melanogaster

<220>
<221> misc_feature
<223> GenBank ID No: g116443

<400> 38
Met Ala Ser Val Ala Ala Trp Leu Pro Phe Ala Arg Ala Ala Ala
1 5 10 15
Ile Gly Trp Val Pro Ile Ala Thr His Pro Leu Pro Pro Pro Pro
20 25 30
Met Pro Lys Asp Arg Arg Lys Thr Asp Asp Glu Lys Leu Leu Ile
35 40 45
Asn Val Ser Gly Arg Arg Phe Glu Thr Trp Arg Asn Thr Leu Glu
50 55 60
Lys Tyr Pro Asp Thr Leu Leu Gly Ser Asn Glu Arg Glu Phe Phe
65 70 75
Tyr Asp Glu Asp Cys Lys Glu Tyr Phe Phe Asp Arg Asp Pro Asp
80 85 90
Ile Phe Arg His Ile Leu Asn Tyr Tyr Arg Thr Gly Lys Leu His
95 100 105
Tyr Pro Lys His Glu Cys Leu Thr Ser Tyr Asp Glu Glu Leu Ala
110 115 120

Phe Phe Gly Ile Met Pro Asp Val Ile Gly Asp Cys Cys Tyr Glu
125 130 135
Asp Tyr Arg Asp Arg Lys Arg Glu Asn Ala Glu Arg Leu Met Asp
140 145 150
Asp Lys Leu Ser Glu Asn Gly Asp Gln Asn Leu Gln Gln Leu Thr
155 160 165
Asn Met Arg Gln Lys Met Trp Arg Ala Phe Glu Asn Pro His Thr
170 175 180
Ser Thr Ser Ala Leu Val Phe Tyr Tyr Val Thr Gly Phe Phe Ile
185 190 195
Ala Val Ser Val Met Ala Asn Val Val Glu Thr Val Pro Cys Gly
200 205 210
His Arg Pro Gly Arg Ala Gly Thr Leu Pro Cys Gly Glu Arg Tyr
215 220 225
Lys Ile Val Phe Phe Cys Leu Asp Thr Ala Cys Val Met Ile Phe
230 235 240
Thr Ala Glu Tyr Leu Leu Arg Leu Phe Ala Ala Pro Asp Arg Cys
245 250 255
Lys Phe Val Arg Ser Val Met Ser Ile Ile Asp Val Val Ala Ile
260 265 270
Met Pro Tyr Tyr Ile Gly Leu Gly Ile Thr Asp Asn Asp Asp Val
275 280 285
Ser Gly Ala Phe Val Thr Leu Arg Val Phe Arg Val Phe Arg Ile
290 295 300
Phe Lys Phe Ser Arg His Ser Gln Gly Leu Arg Ile Leu Gly Tyr
305 310 315
Thr Leu Lys Ser Cys Ala Ser Glu Leu Gly Phe Leu Val Phe Ser
320 325 330
Leu Ala Met Ala Ile Ile Ile Phe Ala Thr Val Met Phe Tyr Ala
335 340 345
Glu Lys Asn Val Asn Gly Thr Asn Phe Thr Ser Ile Pro Ala Ala
350 355 360
Phe Trp Tyr Thr Ile Val Thr Met Thr Thr Leu Gly Tyr Gly Asp
365 370 375
Met Val Pro Glu Thr Ile Ala Gly Lys Ile Val Gly Gly Val Cys
380 385 390
Ser Leu Ser Gly Val Leu Val Ile Ala Leu Pro Val Pro Val Ile
395 400 405
Val Ser Asn Phe Ser Arg Ile Tyr His Gln Asn Gln Arg Ala Asp
410 415 420
Lys Arg Lys Ala Gln Arg Lys Ala Arg Leu Ala Arg Ile Arg Ile
425 430 435
Ala Lys Ala Ser Ser Gly Ala Ala Phe Val Ser Lys Lys Lys Ala
440 445 450
Ala Glu Ala Arg Trp Ala Ala Gln Glu Ser Gly Ile Glu Leu Asp
455 460 465
Asp Asn Tyr Arg Asp Glu Asp Ile Phe Glu Leu Gln His His His
470 475 480
Leu Leu Arg Cys Leu Glu Lys Thr Thr Met
485 490

<210> 39
<211> 478
<212> PRT
<213> Polyorchis penicillatus

<220>
<221> misc_feature
<223> GenBank ID No: g1763619

<400> 39
Met Asn Gly Asp Ile Gly Ala Trp Ile Ser Cys Ala Arg Thr Ala
1   5   10   15
Gly Ile Gly Trp Val Pro Ile Ser Ser Lys Glu Pro Ser Ala Tyr
                20   25   30
Leu Asn Lys Gln Val Cys Asn Glu Asn Glu Lys Asn Asn Ala Lys
                35   40   45
Leu Thr Ile Asn Val Ser Gly Arg Arg Tyr Gln Thr Tyr Ser His
                50   55   60
Thr Leu Arg Lys Phe Lys Glu Thr Leu Leu Gly Ser Gln Glu Arg
                65   70   75
Asp Tyr Phe Tyr Asp Glu Ser Leu Glu Glu Tyr Tyr Phe Asp Arg
                80   85   90
Asp Pro Asp Leu Phe Arg His Ile Leu Asn Tyr Tyr Arg Thr Gly
                95   100   105
Lys Leu His Phe Pro Lys Asn Glu Cys Val Ser Ser Phe Glu Asp
                110   115   120
Glu Leu Thr Phe Phe Gly Ile Lys Gly Phe Asn Ile Asn Asn Cys
                125   130   135
Cys Trp Asp Asp Tyr His Asp Lys Lys Arg Glu Cys Thr Glu Arg
                140   145   150
Leu Asn Glu Ser Asp Val Met Leu Thr Ser Ser Glu Ile Asn Glu
                155   160   165
Lys Ser Asp Thr Met Gly Ile Asp Val Gln Met Asn Asn His Gln
                170   175   180
Ala Lys Asn Phe Arg Gln Lys Val His Gly Leu Phe Glu Asn Pro
                185   190   195
Gln Ser Thr Phe Leu Ala Arg Ile Leu Tyr Tyr Ile Thr Gly Phe
                200   205   210
Phe Ile Ala Val Ser Val Gly Ser Thr Ile Ile Glu Thr Ile Asp
                215   220   225
Cys Ser Ala Asn Arg Pro Cys Gly Glu Val Tyr Asn Lys Ile Phe
                230   235   240
Phe Asn Ile Glu Ala Val Cys Val Val Val Phe Thr Ile Glu Tyr
                245   250   255
Leu Ala Arg Leu Tyr Ser Ala Pro Cys Arg Phe Arg His Ala Arg
                260   265   270
Ile Ser Leu Ser Ile Ile Asp Val Ile Ala Ile Leu Pro Phe Tyr
                275   280   285
Ile Gly Leu Ala Met Thr Lys Thr Ser Ile Ser Gly Ala Phe Val
                290   295   300
Ser Leu Arg Val Phe Arg Val Phe Arg Ile Phe Lys Phe Ser Arg
                305   310   315
His Ser Lys Gly Leu Arg Ile Leu Gly Ser Thr Leu Thr Ser Cys
                320   325   330
Ala Ser Glu Leu Gly Phe Leu Leu Phe Ser Leu Ser Met Ala Ile
                335   340   345
Ile Ile Phe Ala Thr Val Val Phe Tyr Val Glu Lys Asp Val Asn
                350   355   360
Asp Ser Asp Phe Thr Ser Ile Pro Ala Ser Phe Trp Tyr Thr Ile
                365   370   375
Val Thr Met Thr Thr Leu Gly Tyr Gly Asp Met Val Pro Lys Thr

```
                380                 385                 390
Ile Pro Gly Lys Leu Val Gly Ser Ile Cys Ser Leu Ser Gly Val
                395                 400                 405
Leu Val Ile Ala Leu Pro Val Pro Val Ile Val Ser Asn Phe Ser
                410                 415                 420
Arg Ile Tyr Leu Gln Asn Gln Arg Ala Asp Lys Arg Arg Ala Asn
                425                 430                 435
Gln Lys Leu Arg Asn Lys Cys Glu Glu Lys Glu Glu Lys Lys Lys
                440                 445                 450
Glu Ser Ser Ser Glu Thr Val Thr Arg Phe Ile Ile Ser Asn Gln
                455                 460                 465
Met Tyr Thr Ile Phe Ser Met Lys Phe Ala Leu Thr Arg
                470                 475


<210> 40
<211> 732
<212> PRT
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> GenBank ID No: g2564072

<400> 40
Met Asp Thr Ser Gly His Phe His Glu Ser Gly Val Gly Asp Leu
  1               5                  10                  15
Asp Glu Asp Pro Lys Cys Pro Cys Pro Ser Ser Gly Asp Glu Gln
                 20                  25                  30
Gln Gln Gln Gln Gln Pro Pro Pro Pro Ser Ala Pro Pro Ala Val
                 35                  40                  45
Pro Gln Gln Pro Pro Gly Pro Leu Leu Gln Pro Gln Pro Pro Gln
                 50                  55                  60
Leu Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
                 65                  70                  75
Gln Gln Gln Gln Gln Ala Pro Leu His Pro Leu Pro Gln Leu Ala
                 80                  85                  90
Gln Leu Gln Ser Gln Val Val His Pro Gly Leu Leu His Ser Ser
                 95                 100                 105
Pro Thr Ala Phe Arg Ala Pro Asn Ser Ala Asn Ser Thr Ala Ile
                110                 115                 120
Leu His Pro Ser Ser Arg Gln Gly Ser Gln Leu Asn Leu Asn Asp
                125                 130                 135
His Leu Val Gly His Ser Pro Ser Ser Thr Ala Thr Ser Gly Pro
                140                 145                 150
Gly Gly Gly Ser Arg His Arg Gln Ala Ser Pro Val Val His Arg
                155                 160                 165
Arg Asp Ser Asn Pro Phe Thr Glu Ile Ala Met Ser Ser Cys Lys
                170                 175                 180
Tyr Ser Gly Gly Val Met Lys Pro Leu Asn Arg Leu Ser Ala Ser
                185                 190                 195
Arg Arg Asn Leu Ile Glu Ala Glu Pro Glu Gly Gln Pro Leu Gln
                200                 205                 210
Leu Phe Ser Pro Ser Asn Pro Pro Glu Ile Ile Ile Ser Ser Arg
                215                 220                 225
Glu Asp Asn His Ala His Gln Thr Leu Leu His His Pro Asn Ala
```

```
                230                 235                 240
Thr His Asn His Gln His Ala Gly Thr Thr Ala Gly Ser Thr Thr
                245                 250                 255
Phe Pro Lys Ala Asn Lys Arg Lys Asn Gln Asn Ile Gly Tyr Lys
                260                 265                 270
Leu Gly His Arg Arg Ala Leu Phe Glu Lys Arg Lys Arg Leu Ser
                275                 280                 285
Asp Tyr Ala Leu Ile Phe Gly Met Phe Gly Ile Val Val Met Val
                290                 295                 300
Ile Glu Thr Glu Leu Ser Trp Gly Leu Tyr Ser Lys Asp Ser Met
                305                 310                 315
Phe Ser Leu Ala Leu Lys Cys Leu Ile Ser Leu Ser Thr Ile Ile
                320                 325                 330
Leu Leu Gly Leu Ile Ile Ala Tyr His Thr Arg Glu Val Gln Leu
                335                 340                 345
Phe Val Ile Asp Asn Gly Ala Asp Asp Trp Arg Ile Ala Met Thr
                350                 355                 360
Tyr Glu Arg Ile Leu Tyr Ile Ser Leu Glu Met Leu Val Cys Ala
                365                 370                 375
Ile His Pro Ile Pro Gly Glu Tyr Lys Phe Phe Trp Thr Ala Arg
                380                 385                 390
Leu Ala Phe Ser Tyr Thr Pro Ser Arg Ala Glu Ala Asp Val Asp
                395                 400                 405
Ile Ile Leu Ser Ile Pro Met Phe Leu Arg Leu Tyr Leu Ile Ala
                410                 415                 420
Arg Val Met Leu Leu His Ser Lys Leu Phe Thr Asp Ala Ser Ser
                425                 430                 435
Arg Ser Ile Gly Ala Leu Asn Lys Ile Asn Phe Asn Thr Arg Phe
                440                 445                 450
Val Met Lys Thr Leu Met Thr Ile Cys Pro Gly Thr Val Leu Leu
                455                 460                 465
Val Phe Ser Ile Ser Leu Trp Ile Ile Ala Ala Trp Thr Val Arg
                470                 475                 480
Val Cys Glu Arg Tyr His Asp Gln Gln Asp Val Thr Ser Asn Phe
                485                 490                 495
Leu Gly Ala Met Trp Leu Ile Ser Ile Thr Phe Leu Ser Ile Gly
                500                 505                 510
Tyr Gly Asp Met Val Pro His Thr Tyr Cys Gly Lys Gly Val Cys
                515                 520                 525
Leu Leu Thr Gly Ile Met Gly Ala Gly Cys Thr Ala Leu Val Val
                530                 535                 540
Ala Val Val Ala Arg Lys Leu Glu Leu Thr Lys Ala Glu Lys His
                545                 550                 555
Val His Asn Phe Met Met Asp Thr Gln Leu Thr Lys Arg Ile Lys
                560                 565                 570
Asn Ala Ala Ala Asn Val Leu Arg Glu Thr Trp Leu Ile Tyr Lys
                575                 580                 585
His Thr Lys Leu Leu Lys Lys Ile Asp His Ala Lys Val Arg Lys
                590                 595                 600
His Gln Arg Lys Phe Leu Gln Ala Ile His Gln Leu Arg Gly Val
                605                 610                 615
Lys Met Glu Gln Arg Lys Leu Ser Asp Gln Ala Asn Thr Leu Val
                620                 625                 630
Asp Leu Ser Lys Met Gln Asn Val Met Tyr Asp Leu Ile Thr Glu
                635                 640                 645
Leu Asn Asp Arg Ser Glu Asp Leu Glu Lys Gln Ile Gly Ser Leu
```

650 655 660
Glu Ser Lys Leu Glu His Leu Thr Ala Ser Phe Asn Ser Leu Pro
665 670 675
Leu Leu Ile Ala Asp Thr Leu Arg Gln Gln Gln Gln Gln Leu Leu
680 685 690
Thr Ala Phe Val Glu Ala Arg Gly Ile Ser Val Ala Val Gly Thr
695 700 705
Ser His Ala Pro Pro Ser Asp Ser Pro Ile Gly Ile Ser Ser Thr
710 715 720
Ser Phe Pro Thr Pro Tyr Thr Ser Ser Ser Ser Cys
725 730

<210> 41
<211> 269
<212> PRT
<213> Rattus norvegicus

<220>
<221> misc_feature
<223> GenBank ID No: g2350843

<400> 41
Met Ala Gly Ser Val Leu Glu Asn Ile Gln Ser Val Leu Gln Lys
1 5 10 15
Thr Trp Val Arg Glu Phe Leu Ala Glu Phe Leu Asn Thr Tyr Val
20 25 30
Leu Met Val Phe Gly Leu Gly Ser Val Ala His Met Val Leu Gly
35 40 45
Glu Arg Leu Gly Ser Tyr Leu Gly Val Asn Leu Gly Phe Gly Phe
50 55 60
Gly Val Thr Met Gly Ile His Val Ala Gly Gly Ile Ser Gly Ala
65 70 75
His Met Asn Pro Ala Val Thr Phe Thr Asn Cys Ala Leu Gly Arg
80 85 90
Met Ala Gly Arg Lys Phe Pro Ile Tyr Val Leu Gly Gln Phe Leu
95 100 105
Gly Ser Phe Leu Ala Ala Ala Thr Thr Tyr Leu Ile Phe Tyr Gly
110 115 120
Ala Ile Asn His Tyr Ala Gly Glu Thr Leu Leu Val Thr Gly Pro
125 130 135
Lys Ser Thr Ala Asn Ile Phe Ala Thr Tyr Leu Pro Glu His Met
140 145 150
Thr Leu Trp Arg Gly Phe Val Asp Glu Val Phe Val Thr Gly Met
155 160 165
Leu Gln Leu Cys Ile Phe Ala Ile Thr Asp Lys Leu Asn Ser Pro
170 175 180
Ala Leu Gln Gly Thr Glu Pro Leu Met Ile Gly Ile Leu Val Cys
185 190 195
Val Leu Gly Val Ser Leu Gly Met Asn Thr Gly Tyr Ala Ile Asn
200 205 210
Pro Ser Arg Asp Leu Pro Pro Arg Phe Phe Thr Phe Ile Ala Gly
215 220 225
Trp Gly Lys Lys Val Phe Ser Ala Gly Asn Asn Trp Trp Trp Val
230 235 240
Pro Val Val Ala Pro Leu Leu Gly Ala Tyr Leu Gly Gly Ile Val

```
                245                 250                 255
Tyr Leu Gly Leu Ile His Ala Gly Ile Pro Pro Gln Gly Ser
                260                 265
```

<210> 42
<211> 266
<212> PRT
<213> Mus musculus

<220>
<221> misc_feature
<223> GenBank ID No: g192647

<400> 42

```
Met Asn Trp Gly Phe Leu Gln Gly Ile Leu Ser Gly Val Asn Lys
  1             5                  10                  15
Tyr Ser Thr Ala Leu Gly Arg Ile Trp Leu Ser Val Val Phe Ile
               20                  25                  30
Phe Arg Val Leu Val Tyr Val Val Ala Ala Glu Glu Val Trp Asp
               35                  40                  45
Asp Asp Gln Lys Asp Phe Ile Cys Asn Thr Lys Gln Pro Gly Cys
               50                  55                  60
Pro Asn Val Cys Tyr Asp Glu Phe Phe Pro Val Ser His Val Arg
               65                  70                  75
Leu Trp Ala Leu Gln Leu Ile Leu Val Thr Cys Pro Ser Leu Leu
               80                  85                  90
Val Val Met His Val Ala Tyr Arg Glu Glu Arg Glu Arg Lys His
               95                 100                 105
Arg Leu Lys His Gly Pro Asn Ala Pro Ala Leu Tyr Ser Asn Leu
              110                 115                 120
Ser Lys Lys Arg Gly Gly Leu Trp Trp Thr Tyr Leu Leu Ser Leu
              125                 130                 135
Ile Phe Lys Ala Ala Val Asp Ser Gly Phe Leu Tyr Ile Phe His
              140                 145                 150
Cys Ile Tyr Lys Asp Tyr Asp Met Pro Arg Val Val Ala Cys Ser
              155                 160                 165
Val Thr Pro Cys Pro His Thr Val Asp Cys Tyr Ile Ala Arg Pro
              170                 175                 180
Thr Glu Lys Lys Val Phe Thr Tyr Phe Met Val Val Thr Ala Ala
              185                 190                 195
Ile Cys Ile Leu Leu Asn Leu Ser Glu Val Val Tyr Leu Val Gly
              200                 205                 210
Lys Arg Cys Met Glu Val Phe Arg Pro Arg Arg Arg Lys Ala Ser
              215                 220                 225
Arg Arg His Gln Leu Pro Asp Thr Cys Pro Pro Tyr Val Ile Ser
              230                 235                 240
Lys Gly Gly His Pro Gln Asp Glu Ser Val Ile Leu Thr Lys Ala
              245                 250                 255
Gly Met Ala Thr Val Asp Ala Gly Val Tyr Pro
              260                 265
```

<210> 43
<211> 191
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> GenBank ID No: g1055345

<400> 43
Met Val Lys Lys Leu Val Met Ala Gln Lys Arg Gly Glu Thr Arg
1 5 10 15
Ala Leu Cys Leu Gly Val Thr Met Val Val Cys Ala Val Ile Thr
20 25 30
Tyr Tyr Ile Leu Val Thr Thr Val Leu Pro Leu Tyr Gln Lys Ser
35 40 45
Val Trp Thr Gln Glu Ser Lys Cys His Leu Ile Glu Thr Asn Ile
50 55 60
Arg Asp Gln Glu Glu Leu Lys Gly Lys Lys Val Pro Gln Tyr Pro
65 70 75
Cys Leu Trp Val Asn Val Ser Ala Ala Gly Arg Trp Ala Val Leu
80 85 90
Tyr His Thr Glu Asp Thr Arg Asp Gln Asn Gln Gln Cys Ser Tyr
95 100 105
Ile Pro Gly Ser Val Asp Asn Tyr Gln Thr Ala Arg Ala Asp Val
110 115 120
Glu Lys Val Arg Ala Lys Phe Gln Glu Gln Gln Val Phe Tyr Cys
125 130 135
Phe Ser Ala Pro Arg Gly Asn Glu Thr Ser Val Leu Phe Gln Arg
140 145 150
Leu Tyr Gly Pro Gln Ala Leu Leu Phe Ser Leu Phe Trp Pro Thr
155 160 165
Phe Leu Leu Thr Gly Gly Leu Leu Ile Ile Ala Met Val Lys Ser
170 175 180
Asn Gln Tyr Leu Ser Ile Leu Ala Ala Gln Lys
185 190

<210> 44
<211> 308
<212> PRT
<213> Caenorhabditis elegans

<220>
<221> misc_feature
<223> GenBank ID No: g3292929

<400> 44
Met Ser Thr Val Phe Ile Asn Ser Arg Lys Ser Pro Asn Val Leu
1 5 10 15
Lys Lys Gln Gly Thr Asp Gln Trp Val Lys Leu Asn Val Gly Gly
20 25 30
Thr Tyr Phe Leu Thr Thr Lys Thr Thr Leu Ser Arg Asp Pro Asn
35 40 45
Ser Phe Leu Ser Arg Leu Ile Gln Glu Asp Cys Asp Leu Ile Ser
50 55 60
Asp Arg Asp Glu Thr Gly Ala Tyr Leu Ile Asp Arg Asp Pro Lys
65 70 75
Tyr Phe Ala Pro Val Leu Asn Tyr Leu Arg His Gly Lys Leu Val

80 85 90
Leu Asp Gly Val Ser Glu Glu Gly Val Leu Glu Glu Ala Glu Phe
95 100 105
Tyr Asn Val Thr Gln Leu Ile Ala Leu Leu Lys Glu Cys Ile Leu
110 115 120
His Arg Asp Gln Arg Pro Gln Thr Asp Lys Lys Arg Val Tyr Arg
125 130 135
Val Leu Gln Cys Arg Glu Gln Glu Leu Thr Gln Met Ile Ser Thr
140 145 150
Leu Ser Asp Gly Trp Arg Phe Glu Gln Leu Ile Ser Met Gln Tyr
155 160 165
Thr Asn Tyr Gly Pro Phe Glu Asn Asn Glu Phe Leu Cys Val Val
170 175 180
Ser Lys Glu Cys Gly Thr Thr Ala Gly Arg Glu Leu Glu Leu Asn
185 190 195
Asp Arg Ala Lys Val Leu Gln Gln Lys Gly Ser Arg Ile Asn Thr
200 205 210
Ile Ser His Ser Ala Thr Pro Thr Gln His Gln Leu Asp Ala Ala
215 220 225
Lys Glu Ala Arg Ala Thr Ala Thr Ala Thr Ser Asn Thr Thr Asn
230 235 240
His Thr Arg Ser Asp Gln Thr Gln Pro Gln Ala Gln Ile Thr His
245 250 255
Gln Asp Gln Pro Glu Ser Pro Lys Gln Ser Pro Gln Gly Asp Tyr
260 265 270
Ala Ser Phe Ala Phe Glu Thr Lys Leu Thr Gly Thr Thr Ala Ile
275 280 285
Arg Phe Ser Pro Leu Trp Pro Phe Cys Ala Leu Tyr Glu Val Cys
290 295 300
Ala Gly Val His Val Phe Asn Leu
305

<210> 45
<211> 295
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> GenBank ID No: g2887407

<400> 45
Met Gln Pro Glu Gly Ala Glu Lys Gly Lys Ser Phe Lys Gln Arg
1 5 10 15
Leu Val Leu Lys Ser Ser Leu Ala Lys Glu Thr Leu Ser Glu Phe
20 25 30
Leu Gly Thr Phe Ile Leu Ile Val Leu Gly Cys Gly Cys Val Ala
35 40 45
Gln Ala Ile Leu Ser Arg Gly Arg Phe Gly Gly Val Ile Thr Ile
50 55 60
Asn Val Gly Phe Ser Met Ala Val Ala Met Ala Ile Tyr Val Ala
65 70 75
Gly Gly Val Ser Gly Gly His Ile Asn Pro Ala Val Ser Leu Ala
80 85 90
Met Cys Leu Phe Gly Arg Met Lys Trp Phe Lys Leu Pro Phe Tyr

```
                95                  100                 105
Val Gly Ala Gln Phe Leu Gly Ala Phe Val Gly Ala Ala Thr Val
                110                 115                 120
Phe Gly Ile Tyr Tyr Asp Gly Leu Met Ser Phe Ala Gly Gly Lys
                125                 130                 135
Leu Leu Ile Val Gly Glu Asn Ala Thr Ala His Ile Phe Ala Thr
                140                 145                 150
Tyr Pro Ala Pro Tyr Leu Ser Leu Ala Asn Ala Phe Ala Asp Gln
                155                 160                 165
Val Val Ala Thr Met Ile Leu Leu Ile Ile Val Phe Ala Ile Phe
                170                 175                 180
Asp Ser Arg Asn Leu Gly Ala Pro Arg Gly Leu Glu Pro Ile Ala
                185                 190                 195
Ile Gly Leu Leu Ile Ile Val Ile Ala Ser Ser Leu Gly Leu Asn
                200                 205                 210
Ser Gly Cys Ala Met Asn Pro Ala Arg Asp Leu Ser Pro Arg Leu
                215                 220                 225
Phe Thr Ala Leu Ala Gly Trp Gly Phe Glu Val Phe Arg Ala Gly
                230                 235                 240
Asn Asn Phe Trp Trp Ile Pro Val Val Gly Pro Leu Val Gly Ala
                245                 250                 255
Val Ile Gly Gly Leu Ile Tyr Val Leu Val Ile Glu Ile His His
                260                 265                 270
Pro Glu Pro Asp Ser Val Phe Lys Ala Glu Gln Ser Glu Asp Lys
                275                 280                 285
Pro Glu Lys Tyr Glu Leu Ser Val Ile Met
                290                 295
```